



UNIVERSIDADE ESTADUAL DE CAMPINAS

Instituto de Matemática, Estatística
e Computação Científica

CLAUDIA AKEMY KODA

MODELOS MISTOS PARA RESPOSTAS POSITIVAS
AUMENTADAS EM ZERO

CAMPINAS
2016

CLAUDIA AKEMY KODA

**MODELOS MISTOS PARA RESPOSTAS POSITIVAS
AUMENTADAS EM ZERO**

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestra em Estatística.

Orientadora: Mariana Rodrigues Motta

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO DEFENDIDA PELA ALUNA CLAUDIA AKEMY KODA, E ORIENTADA PELA PROFA. DRA. MARIANA RODRIGUES MOTTA.

Assinatura da Orientadora

Mariana Rodrigues Motta

**CAMPINAS
2016**

Agência(s) de fomento e nº(s) de processo(s): CAPES

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Ana Regina Machado - CRB 8/5467

K816m Koda, Claudia Akemy, 1990-
Modelos mistos para respostas positivas aumentadas em zero / Claudia Akemy Koda. – Campinas, SP : [s.n.], 2016.

Orientador: Mariana Rodrigues Motta.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Modelos mistos. 2. Inferência bayesiana. 3. Métodos MCMC. 4. Distribuição (Probabilidades). I. Motta, Mariana Rodrigues, 1975-. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Mixed models for zero-augmented positive responses

Palavras-chave em inglês:

Mixed models

Bayesian inference

MCMC methods

Distribution (Probability theory)

Área de concentração: Estatística

Titulação: Mestra em Estatística

Banca examinadora:

Mariana Rodrigues Motta [Orientador]

Viviana Giampaoli

Filidor Edilfonso Vilca Labra

Data de defesa: 29-02-2016

Programa de Pós-Graduação: Estatística

Dissertação de Mestrado defendida em 29 de fevereiro de 2016 e aprovada

Pela Banca Examinadora composta pelos Profs. Drs.

Prof(a). Dr(a). MARIANA RODRIGUES MOTTA

Prof(a). Dr(a). VIVIANA GIAMPAOLI

Prof(a). Dr(a). FILIDOR EDILFONSO VILCA LABRA

A Ata da defesa com as respectivas assinaturas dos membros encontra-se no processo de vida acadêmica do aluno.

Para minha família.

Agradecimentos

Agradeço à minha família por todo o incentivo, apoio e compreensão durante esse período de estudos. Em especial, aos meus pais e irmã pois tenho certeza que esse trabalho não poderia ser realizado sem sua ajuda todos os dias.

Aos meus queridos amigos Jacqueline, Fabíola, Tomás, Paula, Diego, Paulo, Gustavo e Andrea por sua amizade e por entenderem minha ausência em muitos momentos.

Às minhas amigas Bárbara e Mariana pelas longas conversas e por estarem sempre ao meu lado.

Aos meus amigos e colegas de mestrado Lígia e Gustavo, por todo o incentivo nos momentos difíceis e pelos bons momentos que compartilhamos. Também ao Chico, pela companhia nos muitos finais de semana resolvendo listas e trabalhos. À Monique pelos momentos de distração. Aos queridos Gabriel e Darcy por todos os conselhos. À todos os outros colegas por seu companheirismo e apoio.

Agradeço muito à minha orientadora, Professora Mariana Rodrigues Motta, por todo o seu apoio, paciência e disponibilidade ao longo de todo o trabalho.

Aos Professores Filidor Edilfonso Vilca Labra e Viviana Giampaoli por aceitarem fazer parte da banca examinadora e por suas sugestões que contribuíram para este trabalho.

Ao Professor Eliseu Verly Junior, por sua grande contribuição ao disponibilizar os dados para a aplicação e por sua disponibilidade em responder minhas dúvidas.

Aos professores do Departamento de Estatística do IMECC por contribuírem para minha formação.

À CAPES, pelo apoio financeiro.

Resumo

Em geral, a distribuição de probabilidade de respostas positivas é assimétrica, afastando a suposição de normalidade. Algumas transformações para se obter normalidade dos dados podem não ser apropriadas quando existe a presença do valor zero, como é o caso da transformação logaritmo. Uma alternativa é analisar os dados sem as observações iguais a zero. No entanto, isso causaria superestimação da média. Neste trabalho, estudamos uma classe de modelos mistos para respostas que assumem valores positivos e o valor zero a partir de um modelo de duas partes, combinando uma distribuição de probabilidade contínua para os valores positivos a um ponto de massa em zero. Utilizamos distribuições positivas pertencentes à família exponencial de dois parâmetros parametrizadas através da média para modelar respostas positivas. Adicionalmente, modelamos o logito da probabilidade do zero ocorrer e o logaritmo da média dos valores positivos através de covariáveis, induzindo uma correlação entre as duas partes através de efeitos aleatórios correlacionados. Os efeitos aleatórios acomodam a variabilidade entre as diferentes unidades experimentais e a variabilidade dentro uma mesma unidade experimental quando houver medidas repetidas. Aspectos inferenciais do modelo são estudados a partir de uma abordagem Bayesiana. Realizamos um estudo de simulação para acessar a qualidade das estimativas a partir de quatro tipos de distribuições a priori para as componentes de dispersão dos efeitos aleatórios. Um outro estudo de simulação foi realizado para acessar a qualidade das estimativas das componentes de variância dos efeitos aleatórios a partir de diferentes combinações do tamanho amostral e número de medidas repetidas para cada unidade experimental. Finalmente, o modelo apresentado foi utilizado para modelar dados sobre consumo habitual.

Palavras-chave: Modelos mistos. Inferência Bayesiana. Métodos MCMC. Distribuição (Probabilidades).

Abstract

In general, the probability distribution of positive responses is asymmetrical, invalidating the normality assumption. Some transformations to obtain normality of the data may not be appropriate when there is the presence of zero, as it is the case of the logarithm transformation. An alternative to analyze the data is to discard observations equal to zero. However, this would cause overestimation of the mean. Here, we study a class of mixed models for responses that can be either positive or zero through a model of two parts, combining a continuous probability distribution for the positive values to a point mass at zero. We consider continuous probability distributions in the class of exponential family distributions with two parameters parameterized by the mean. Additionally, we model the logit probability of the zero and the logarithm of the mean of positive values by means of covariates, inducing a correlation between them through correlated random effects. Random effects accommodate variability among different experimental units and the variability within the same experimental unit when there are repeated measurements. Inferential aspects of model are studied from a Bayesian perspective. We conducted a simulation study to access the quality of the estimates from four types of prior distributions for the dispersion of components of random effects. Another simulation study was performed to access the quality of the estimates of the variance components of random effects from different combinations of sample size and number of repeated measures for each experimental unit. Finally, the model was used to model data on usual intake.

Keywords: Mixed models. Bayesian Inference. MCMC methods. Distribution (Probability theory).

Lista de Ilustrações

1.1	Função densidade $\text{Gama}(\mu, \phi)$ parametrizada com $\mu = 2$ em (a) e $\phi = 5$ em (b).	20
1.2	Função densidade $\text{IG}(\mu, \phi)$ com $\mu = 2$ em (a) e $\phi = 5$ em (b).	22
1.3	Função densidade $\text{LN}(\mu, \phi)$ com $\mu = 2$ em (a) e $\phi = 5$ em (b).	25
4.1	Distribuição do consumo positivo de FLV ao longo das visitas.	57
4.2	Frequência de dias sem consumo de FLV para os indivíduos do estudo.	57
4.3	Intervalos de credibilidade de 95% simétricos para os parâmetros dos efeitos fixos para a probabilidade de valor zero (esquerda) e para a média dos valores positivos (direita) no modelo com a distribuição Gama.	59
4.4	Intervalos de credibilidade de 95% simétricos para os parâmetros dos efeitos fixos para a probabilidade de valor zero (esquerda) e para a média dos valores positivos (direita) no modelo com a distribuição Inversa Gaussiana.	59
4.5	Intervalos de credibilidade de 95% simétricos para os parâmetros dos efeitos fixos para a probabilidade de valor zero (esquerda) e para a média dos valores positivos (direita) no modelo com a distribuição Lognormal.	60
4.6	Medidas de divergência de Kullback-Liebler, J-distância e L_1 -distância para o modelo Gama.	62
4.7	Medidas de divergência de Kullback-Liebler, J-distância e L_1 -distância para o modelo Inversa Gaussiana.	63
4.8	Medidas de divergência de Kullback-Liebler, J-distância e L_1 -distância para o modelo Lognormal.	63
4.9	Histograma da densidade preditiva do Consumo positivo de FLV com as curvas para a densidade preditiva a posteriori gerada dos modelos Gama, Inversa Gaussiana e Lognormal sobrepostas.	65
4.10	Box-plot das médias preditas a posteriori (em preto) e dos valores positivos observados (em cinza).	66
C.1	Gráfico do tipo Trace plot para os parâmetros associados aos efeitos fixos da probabilidade de zero do modelo final utilizando a distribuição Gama.	87
C.2	Gráfico do tipo Trace plot para os parâmetros associados aos efeitos fixos da média dos valores positivos do modelo final utilizando a distribuição Gama.	88
C.3	Gráfico do tipo Trace plot para os parâmetros associados aos efeitos fixos das componentes de variância do modelo final utilizando a distribuição Gama.	88

Lista de Tabelas

4.1	Descrição das variáveis presentes no estudo.	55
4.2	Porcentagem de zeros e medidas-resumo para os valores positivos do consumo de FLV em cada visita.	56
4.3	Estimativas a posteriori dos parâmetros e critérios de comparação para os modelos finais utilizando as distribuições Gama, Inversa Gaussiana e Lognormal. . .	61
A.1	Resultados da simulação segundo ER, EQM e PC para $\rho = 0$	74
A.2	Resultados da simulação segundo ER, EQM e PC para $\rho = 0,5$	75
A.3	Resultados da simulação segundo ER, EQM e PC para $\rho = 0,1$	76
A.4	Resultados da simulação segundo ER, EQM e PC para $\rho = 0,9$	77
A.5	Resultados da simulação segundo ER, EQM e PC para $\rho = -0,5$	78
A.6	Resultados da simulação segundo ER, EQM e PC para $\rho = -0,9$	79
B.1	Resultados da simulação segundo ER, EQM e PC para $\rho = 0,5$	81
B.2	Resultados da simulação segundo ER, EQM e PC para $\rho = 0,1$	82
B.3	Resultados da simulação segundo ER, EQM e PC para $\rho = 0,9$	83
B.4	Resultados da simulação segundo ER, EQM e PC para $\rho = -0,5$	84
B.5	Resultados da simulação segundo ER, EQM e PC para $\rho = -0,9$	85
C.1	Valores \hat{R} para os parâmetros do modelo final utilizando as distribuições Gama, Inversa Gaussiana e Lognormal.	86

Sumário

1	Introdução	12
1.1	Modelos para respostas positivas aumentadas em zero	15
1.1.1	O modelo	15
1.1.2	Família exponencial de dois parâmetros	16
1.1.3	Modelo de regressão para respostas positivas aumentadas em zero com efeitos aleatórios	25
1.1.4	Verossimilhança para respostas positivas aumentadas em zero	28
2	Método Inferencial	29
2.1	Inferência Bayesiana	30
2.1.1	O amostrador de Gibbs	32
2.2	Análise Bayesiana do modelo de regressão para respostas positivas aumentadas de zero	33
2.3	Critérios de comparação de modelos	43
2.4	Análise de observações influentes	46
3	Estudo de simulação sobre a distribuição a priori das componentes de dispersão dos efeitos aleatórios	48
3.1	Introdução	48
3.2	Estudo de simulação da distribuição a priori das componentes de variância a partir do modelo de regressão Gama aumentada em zero	50
3.3	Estudo de simulação a partir do número de observações por unidade amostral para o modelo Gama aumentada em zero	53
4	Aplicação	54
4.1	Análise descritiva	55
4.2	Análise Inferencial	56
5	Considerações finais	67
	Referências	69
A	Estudo de simulação para as priors	73
B	Estudo de simulação para o número de observações por indivíduo	80
C	Análise da convergência	86

Capítulo 1

Introdução

A distribuição de uma variável aleatória semicontínua combina uma distribuição de probabilidade contínua, em geral assimétrica, a um ponto de massa em uma ou mais localidades (Olsen e Schafer, 2001). Em estudos aplicados à área biológica, muitas vezes temos que a resposta de interesse assume valores positivos num intervalo contínuo e o valor zero. Para esse tipo de variável atribui-se uma distribuição de probabilidade contínua combinada a um ponto de massa em zero. Um exemplo ocorre nos estudos de consumo habitual de alimentos e nutrientes (Carriquiry, 2003). Outros exemplos deste tipo de variável aparecem em estudos sobre os gastos com saúde (Zhang et al., 2006, Tooze et al., 2002) e atividade física (Lee et al., 2010).

Em geral, a distribuição de probabilidade de respostas positivas é assimétrica (à direita), afastando a suposição de normalidade para ajustar esse tipo de dado. Algumas transformações para se obter normalidade dos dados podem não ser apropriadas quando existe a presença do valor zero, como é o caso da função logaritmo. Uma alternativa é analisar os dados sem as observações iguais a zero. No entanto, isso causaria uma superestimação da média (Rodrigues-Motta et al., 2015).

Modelos com inflação de zeros são populares por acomodar dados com excesso de zeros (Lambert, 1992). Esses modelos tem a forma de uma mistura de duas componentes e cada uma delas acomoda o zero; uma componente acomoda o zero através de um ponto de massa em zero e outra através de um modelo de contagem. No primeiro caso, o zero é chamado de “zero perfeito” e no segundo caso é tido como zero ao acaso. No entanto, para dados positivos aumentados de zero, o zero aparece somente na componente degenerada em zero, enquanto que a outra componente contém uma distribuição no suporte $(0, \infty)$. Por essa razão, neste estudo adota-se o termo aumentado em contrapartida a inflacionado.

Um outro exemplo de modelos aumentados ocorre quando o interesse é modelar uma proporção. Ospina e Ferrari (2012) propõem um modelo de regressão Beta aumentada em zero ou

em um. Galvis et al. (2014) apresentam um modelo de regressão Beta aumentada em zero e em um com efeitos aleatórios analisando aspectos inferenciais sob a metodologia Bayesiana.

Baseado em Rodrigues-Motta et al. (2015), o objetivo principal é estudar modelos para respostas positivas aumentadas em zero, utilizando distribuições positivas da família exponencial de dois parâmetros, via discussão teórica e prática por meio da análise de um conjunto de dados sobre consumo habitual. O modelo proposto é um modelo de duas partes (Duan et al., 1983), onde uma parte acomoda o valor zero e a outra parte acomoda a resposta positiva, de modo que a probabilidade de observar uma resposta igual a zero e a média dos valores positivos são modelados separadamente, através de covariáveis que podem afetar os dois processos de maneiras distintas. Efeitos aleatórios associados a cada unidade experimental são incorporados à essa estrutura, acomodando medidas repetidas. Exemplos de modelos de duas partes com efeitos aleatórios encontram-se em Olsen e Schafer (2001), Tooze et al. (2002) e Liu et al. (2008)

Para dados de consumo habitual, Kipnis et al. (2009) e Tooze et al. (2006) utilizam modelos de duas partes com efeitos aleatórios. A primeira parte modela a probabilidade de consumo positivo e a segunda modela diretamente o consumo positivo a partir de transformação Box-Cox dos dados, para que a distribuição dos dados transformados possa ser aproximada pela distribuição normal. Porém, utilizando transformação Box-Cox, a transformação dos dados para sua escala original pode não ter forma exata, sendo necessário utilizar aproximações numéricas (Tooze et al., 2010).

Em geral, as recomendações sobre o consumo de nutrientes ou alimentos são baseadas em quantidades diárias. Porém, uma vez que esses nutrientes podem ficar no organismo por mais de um dia e devido à variação no consumo ao longo dos dias, para estudar a adequação às recomendações, deseja-se acessar o consumo habitual, que é definido como o consumo a longo prazo de certo alimento ou nutriente (Dodd et al., 2006).

Para acessar o consumo habitual, comumente se tem utilizado os inquéritos recordatórios de 24 horas, onde os participantes do estudo reportam o consumo alimentar do dia anterior com grande nível de detalhe. Por utilizar dias não consecutivos, reduzem-se problemas com um possível erro sistemático. Ainda assim, o consumo de um indivíduo pode variar muito de um dia a outro (Dodd et al., 2006). Deste modo, para estimar a distribuição do consumo habitual, deve-se levar em conta tanto a variabilidade entre os indivíduos como a entre as observações de um mesmo indivíduo.

Uma vez apresentada a classe de modelos para respostas positivas aumentadas em zero,

estudamos o modelo que, além de acomodar a variabilidade entre unidades experimentais como na aplicação em Rodrigues-Motta et al. (2015), acomoda a variabilidade intra unidade experimental, como por exemplo medidas repetidas. Dentro da classe de distribuições da família exponencial de dois parâmetros, as simulações e a aplicação são feitas a partir das distribuições Gama, Inversa Gaussiana e Lognormal.

A aplicação é feita a partir de um estudo com dados sobre consumo alimentar de indivíduos da cidade do Rio de Janeiro. Estes foram cedidos pelo Prof. Dr. Eliseu Verly Junior da Universidade Estadual do Rio de Janeiro (UERJ) e coordenador do projeto “Variabilidade intrapessoal do consumo dietético e viés em análises de estudos epidemiológicos” com apoio do CNPq (481434/2013-5) e da FAPERJ (E-26/111.822/2012), ao qual agradecemos pela grande contribuição à este trabalho por permitir a análise de dados reais. Os dados foram obtidos entre Março de 2013 a Abril de 2014.

Estudos sobre hábitos alimentares têm sido frequentemente utilizados em estudos epidemiológicos por diversos propósitos, como estimar o consumo de certo alimento ou nutriente para monitorar o estado nutricional de grupos ou indivíduos de uma população e investigar a relação de padrões de dieta com estado de saúde com respeito a certas doenças (Kipnis et al., 2009).

No conjunto de dados, temos 162 indivíduos e para cada indivíduo temos 19 observações sobre o consumo de frutas, legumes e vegetais (em gramas), coletados em entrevistas feitas em dias não consecutivos, além do registro de características sociais e demográficas dos indivíduos. Nesse tipo de estudo, o consumo de certo alimento ou nutriente possui um valor positivo ou possui valor zero, o qual indica que certo alimento ou nutriente não foi consumido nas 24 horas anteriores à entrevista.

Aspectos inferenciais são estudados a partir de uma abordagem Bayesiana e um estudo a partir de quatro prioris para as componentes de variância dos efeitos aleatórios é realizado. Um dos objetivos em pesquisa sobre o consumo habitual é saber qual é o número ideal de entrevistas para se obter boas estimativas da variabilidade intra e entre indivíduos (Carriquiry, 2003). Nesse sentido, apresentamos um estudo de simulação que explora o desempenho do processo de estimação a partir de diferentes quantidades de medidas repetidas e diferentes tamanhos amostrais.

O presente trabalho está organizado da maneira a seguir. Ainda na Introdução apresentamos a classe de modelos para respostas positivas aumentadas em zero e o caso particular a ser usado nesse estudo. O Capítulo 2 apresenta aspectos inferenciais segundo uma abordagem Bayesiana. No Capítulo 3, apresentamos um breve estudo de simulação para a escolha das prioris para as

componentes de variância dos efeitos aleatórios do modelo. No Capítulo 4, apresentamos os resultados da aplicação a partir da metodologia estudada para dados sobre consumo habitual. Por fim, apresentamos uma discussão e considerações finais no Capítulo 5.

1.1 Modelos para respostas positivas aumentadas em zero

Nesta seção, apresentaremos a classe de modelos para respostas positivas aumentadas em zero, onde a distribuição da resposta positiva pertencente à família exponencial de dois parâmetros. Uma vez que, em geral, a distribuição apresenta assimetria, faremos nosso estudo através das distribuições Gama, Lognormal e Inversa Gaussiana (Rodrigues-Motta et al., 2015, Yau et al., 2002).

1.1.1 O modelo

Seja y a realização de uma variável aleatória Y onde $Y \in [0, \infty)$. Uma forma natural de modelar Y é considerar uma mistura de uma densidade $f(\cdot|\boldsymbol{\lambda})$ com uma distribuição degenerada em zero, representada por

$$g(y|\boldsymbol{\lambda}, p) = p \mathbf{I}_{\{y=0\}} + (1-p) f(y|\boldsymbol{\lambda}) \mathbf{I}_{\{y>0\}}, \quad y \geq 0, \quad (1.1.1)$$

onde $p \in (0, 1)$ é o parâmetro de mistura e $\boldsymbol{\lambda} = (\lambda_1, \lambda_2) \in \boldsymbol{\Lambda} \subset \mathbb{R}^2$. Aqui, $\mathbf{I}_{\{\cdot\}}$ representa a função indicadora tal que $\mathbf{I}_{\{\cdot\}}$ é igual a um se a condição $\{\cdot\}$ é satisfeita e zero caso contrário.

Neste trabalho, consideramos $f(\cdot|\boldsymbol{\lambda})$ pertencente à família exponencial com dois parâmetros.

Assim, temos que

$$E[Y] = (1-p) \mu \quad \text{e} \quad Var[Y] = (1-p) \sigma^2 + p (1-p) \mu^2, \quad (1.1.2)$$

onde μ é a esperança e σ^2 é a variância para $Y > 0$ (Aitchison, 1955).

Uma outra forma de escrever (1.1.1) é dada por

$$g(y|\boldsymbol{\lambda}, p) = \left(\frac{p}{1-p} \right)^{\mathbf{I}_{\{y=0\}}} (1-p) f(y|\boldsymbol{\lambda})^{1-\mathbf{I}_{\{y=0\}}}, \quad y \geq 0. \quad (1.1.3)$$

Uma reparametrização de $f(y|\boldsymbol{\lambda})$ é apresentada em termos da média e parâmetro de precisão. Por fim, apresentamos os modelos para respostas positivas aumentadas em zeros com efeitos aleatórios propostos em Rodrigues-Motta et al. (2015).

1.1.2 Família exponencial de dois parâmetros

Seja W uma variável aleatória cuja densidade pertence à família exponencial de dois parâmetros, restrita às variáveis aleatórias positivas, como apresentadas em Bose e Boukai (1993).

Seja

$$f_W(w|\boldsymbol{\lambda}) = h(w) \exp \{ \lambda_1 U_1(w) + \lambda_2 U_2(w) + c(\lambda_1, \lambda_2) \}, \quad (1.1.4)$$

com $w > 0$ e $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$, a função densidade que pertence à família exponencial de dois parâmetros para variáveis aleatórias positivas, com espaço paramétrico definido como

$$\Lambda = \left\{ \boldsymbol{\lambda} = (\lambda_1, \lambda_2) \in \mathbb{R}^2; \exp \{ -c(\lambda_1, \lambda_2) \} = \int h(w) \exp \{ \lambda_1 U_1(w) + \lambda_2 U_2(w) \} dw \right\}. \quad (1.1.5)$$

Seja $\mathbf{U}(W) = (U_1(W), U_2(W))$, como mostrado em Brown (1986)

$$\begin{aligned} E_{\boldsymbol{\lambda}}[\mathbf{U}(W)] &= (\nu_1, \nu_2), & \nu_i &= -\frac{\partial c(\lambda_1, \lambda_2)}{\partial \lambda_i}, & i &= 1, 2, \\ V_{\boldsymbol{\lambda}}[\mathbf{U}(W)] &= (\sigma_{11}, \sigma_{12}, \sigma_{21}, \sigma_{22}), & \sigma_{jk} &= -\frac{\partial^2 c(\lambda_1, \lambda_2)}{\partial \lambda_j \partial \lambda_k}, & j, k &= 1, 2, \end{aligned} \quad (1.1.6)$$

onde $E_{\boldsymbol{\lambda}}[\mathbf{U}(W)]$ é a esperança e $V_{\boldsymbol{\lambda}}[\mathbf{U}(W)]$ é a matriz de covariâncias de $\mathbf{U}(W)$, respectivamente.

Como mostrado em Ferrari e Cribari-Neto (2004) para o modelo de regressão beta, consideramos uma reparametrização das distribuições em (1.1.4) em termos da média e da variância de W , utilizando uma correspondência entre os parâmetros λ_1 e λ_2 com a média μ e o parâmetro de dispersão (ou precisão) ϕ , ou com uma função destes, tal que:

$$f_W(w|\boldsymbol{\lambda}_{\mu}) = h(w) \exp \{ \eta_1(\mu, \phi) U_1(w) + \eta_2(\mu, \phi) U_2(w) + b(\mu, \phi) \}, \quad (1.1.7)$$

com $w > 0$ e $\boldsymbol{\lambda}_{\mu} = (\mu, \phi)$, de modo que $E[W] = \mu$ e $Var[W] = \sigma^2(\mu, \phi)$ é uma função de μ e ϕ .

A seguir, apresentamos duas suposições encontradas em Bar-Lev e Reiser (1982) que nos permitem parametrizar densidades na forma (1.1.7).

Suposição 1.1.1. Seja $\lambda = (\lambda_1, \lambda_2)$. O parâmetro λ_2 pode ser representado como $\lambda_2^* = -\lambda_1 \psi'(\nu_2)$, onde $\psi'(\nu_2) = \partial \psi(\nu_2) / \partial \nu_2$ para uma certa função ψ e ν_2 é a esperança de $U_2(\cdot)$ como definido em (1.1.6).

Suposição 1.1.2. Seja $U_2(w) = a(w)$, onde $a(w)$ é uma função 1 a 1 no suporte de (1.1.4).

As distribuições Normal, Gama e Inversa Gaussiana, assim como densidades resultantes de transformações 1 a 1 destas densidades - por exemplo, a distribuição Lognormal que é obtida através de uma transformação da distribuição Normal - satisfazem as Suposições 1.1.1 e 1.1.2, como mostrado em Bar-Lev e Reiser (1982).

Notamos que para funções densidade na forma (1.1.4) onde $U_2(W) = W$, então $\nu_2 = E[U_2(W)] = E[W]$.

Sob as Suposições 1.1.1 e 1.1.2, Bose e Boukai (1993) mostram como resultado da prova do Lema 3.1 em Bar-Lev e Reiser (1982) que a função $\psi(\nu_2)$ pode ser encontrada através da relação

$$\nu_1 = -\frac{\partial c(\lambda_1, \nu_2)}{\partial \lambda_1} = \psi(\nu_2) + G'(\lambda_1), \quad (1.1.8)$$

onde $G(\lambda_1)$ é uma função infinitamente diferenciável no espaço paramétrico Λ , no qual $G''(\lambda_1) > 0$, para todo $\lambda_1 \in \Lambda$.

Além disso, utilizando a Suposição 1.1.1 e diferenciando $\lambda_2 = -\lambda_1 \psi'(\nu_2)$ em ambos os lados em relação a λ_2 , temos que

$$1 = -\lambda_1 \psi''(\nu_2) \frac{\partial \nu_2}{\partial \lambda_2}, \quad \text{onde} \quad \frac{\partial \nu_2}{\partial \lambda_2} = -\frac{1}{\lambda_1 \psi''(\nu_2)}.$$

Assim, segue que a variância de $U_2(W)$ é dada por:

$$Var[U_2(W)] = \sigma_{22} = -\frac{\partial^2 c(\lambda_1, \lambda_2)}{\partial \lambda_2 \partial \lambda_2} = \frac{\partial \nu_2}{\partial \lambda_2} = -\frac{1}{\lambda_1 \psi''(\nu_2)}. \quad (1.1.9)$$

Utilizando as Suposições 1.1.1 e 1.1.2, representamos a seguir as densidades Gama, Inversa Gaussian e Lognormal a partir da reparametrização proposta em (1.1.7).

(i) **Distribuição Gama**

Seja W uma variável aleatória com distribuição Gama com parâmetros a e b , ou seja, $W \sim \text{Gama}(a, b)$ cuja função densidade é dada por:

$$f_W(w|(a, b)) = \frac{b^a}{\Gamma(a)} w^{a-1} \exp \{-b w\}, \quad w > 0, \quad a > 0, \quad b > 0. \quad (1.1.10)$$

Representando (1.1.10) como em (1.1.4), temos

$$f_W(w|\boldsymbol{\lambda}) = w^{-1} \exp \{\lambda_1 \log(w) + \lambda_2 w + \lambda_1 \log(-\lambda_2) - \log(\Gamma(\lambda_1))\}, \quad w > 0, \quad (1.1.11)$$

tal que $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ com $\lambda_1 = a$ e $\lambda_2 = -b$, $h(w) = w^{-1}$, $U_1(w) = \log(w)$, $U_2(w) = w$ e $c(\lambda_1, \lambda_2) = \lambda_1 \log(-\lambda_2) - \log(\Gamma(\lambda_1))$.

Aqui, $U_2(W) = W$. Com o objetivo de reparametrizar (1.1.11) como (1.1.7), então tomamos $E[W] = \mu$. Considerando a Suposição 1.1.1, então λ_2 pode ser representado como função da média μ , uma vez que $\nu_2 = E[U_2(W)] = E[W] = \mu$.

A partir de (1.1.6) temos que

$$\nu_2 = E[U_2(W)] = -\frac{\partial c(\lambda_1, \lambda_2)}{\partial \lambda_2} = -\frac{\lambda_1}{\lambda_2}. \quad (1.1.12)$$

Usando o resultado (1.1.8), com ν_2 definido em (1.1.12), temos que

$$\nu_1 = E[U_1(W)] = -\frac{\partial c(\lambda_1, \nu_2)}{\partial \lambda_1} = \log(\nu_2) - \log(\lambda_1) + \frac{1}{\Gamma(\lambda_1)} \frac{\partial \Gamma(\lambda_1)}{\partial \lambda_1}, \quad (1.1.13)$$

tal que, $\psi(\nu_2) = \log(\nu_2)$. Assim, temos que $\psi'(\nu_2) = 1/\nu_2$, e $G'(\lambda_1) = (1/\Gamma(\lambda_1)) \times \partial \Gamma(\lambda_1)/\partial \lambda_1 - \log(\lambda_1)$.

Finalmente, seguindo a Suposição 1.1.1, λ_2 pode ser escrito como

$$\lambda_2^* = -\lambda_1 \psi'(\nu_2) = -\frac{\lambda_1}{\nu_2}. \quad (1.1.14)$$

Finalmente,

$$\text{Var}[U_2(W)] = \text{Var}[W] = \frac{\nu_2^2}{\lambda_1}. \quad (1.1.15)$$

Logo, sendo $\nu_2 = \mu$ e fazendo $\lambda_1 = \phi$, temos $(\lambda_1, \lambda_2^*) = (\phi, -\phi/\mu)$ e escrevemos (1.1.11) como

$$\begin{aligned} f_W(w|(\mu, \phi)) &= w^{-1} \exp \left\{ \phi \log(w) - \frac{\phi}{\mu} w + \phi \log \left(\frac{\phi}{\mu} \right) - \log(\Gamma(\phi)) \right\} \\ &= \frac{1}{\Gamma(\phi)} \left(\frac{\phi}{\mu} \right)^\phi w^{\phi-1} \exp \left\{ -\frac{\phi}{\mu} w \right\}, \quad w > 0, \mu > 0, \phi > 0, \end{aligned} \quad (1.1.16)$$

satisfazendo a parametrização (1.1.7), com $h(w) = w^{-1}$, $\eta_1(\mu, \phi) = \phi$, $\eta_2(\mu, \phi) = -\phi/\mu$, $U_1(w) = \log(w)$, $U_2(w) = w$ e $b(\mu, \phi) = \phi \log(\phi/\mu) - \log(\Gamma(\phi))$, tal que

$$E[W] = \mu \quad \text{e} \quad \text{Var}[W] = \frac{\mu^2}{\phi}. \quad (1.1.17)$$

A Figura 1.1 apresenta a função densidade de probabilidade de $W \sim \text{Gama}(\mu, \phi)$ dada por (1.1.16) para alguns valores de μ e ϕ . Em (a) podemos notar que ϕ controla a assimetria da distribuição e que à medida que ϕ aumenta, a distribuição se torna mais concentrada em torno da média, ou seja, a variância diminui. Em (b), notamos que à medida que μ aumenta, a variância também aumenta.

(ii) Distribuição Inversa Gaussiana

Seja W variável aleatória com distribuição Inversa Gaussiana com parâmetros a e b , ou seja, $W \sim \text{IG}(a, b)$ cuja função densidade é dada por:

$$f_W(w|(a, b)) = \sqrt{\frac{a}{2\pi w^3}} \exp \left\{ -\frac{b}{2}w - \frac{a}{2w} + (ab)^{1/2} \right\}, \quad w > 0, a > 0, b \leq 0, \quad (1.1.18)$$

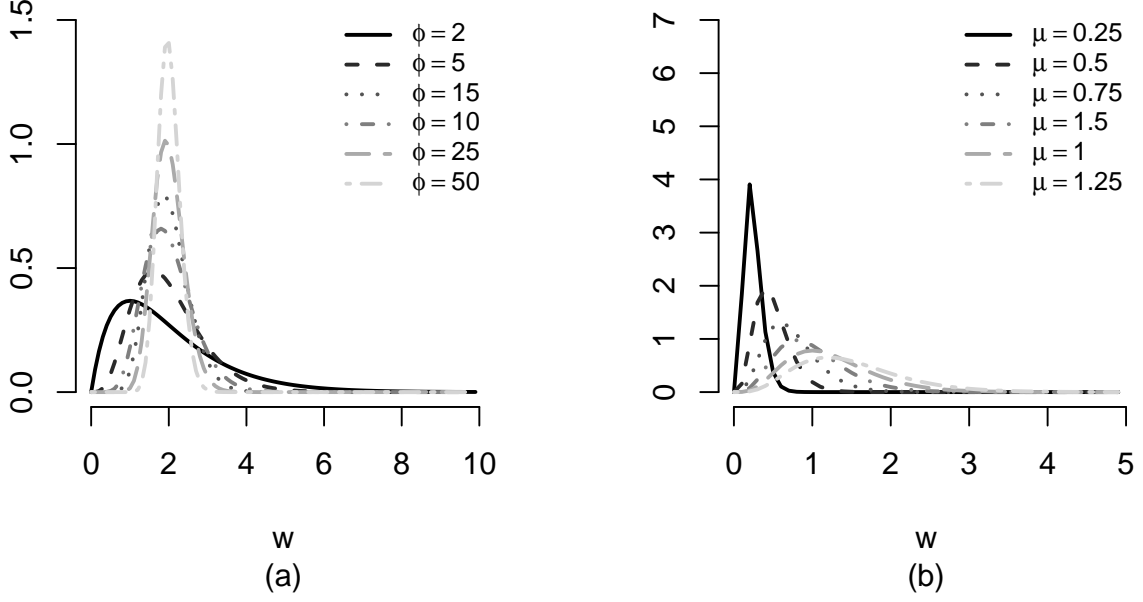


Figura 1.1: Função densidade Gama(μ, ϕ) parametrizada com $\mu = 2$ em (a) e $\phi = 5$ em (b).

Representando (1.1.18) na forma (1.1.4), temos:

$$f_W(w|\boldsymbol{\lambda}) = (2\pi)^{-1/2} w^{-3/2} \exp \left\{ \lambda_1 \frac{1}{w} + \lambda_2 w + 2(\lambda_1 \lambda_2)^{1/2} + \frac{1}{2} \log(-2\lambda_1) \right\}, \quad w > 0, \quad (1.1.19)$$

tal que $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ com $\lambda_1 = -a/2$ e $\lambda_2 = -b/2$, $h(w) = w^{-3/2}$, $U_1(w) = 1/w$, $U_2(w) = w$ e $c(\lambda_1, \lambda_2) = 2(\lambda_1 \lambda_2)^{1/2} + 1/2 \log(-2\lambda_1)$.

Como $U_2(W) = W$, então $\nu_2 = E[U_2(W)] = E[W]$. Além disso, a partir de (1.1.6), temos que

$$\nu_2 = -\frac{\partial c(\lambda_1, \lambda_2)}{\partial \lambda_2} = -\left(\frac{\lambda_1}{\lambda_2}\right)^{1/2}. \quad (1.1.20)$$

Encontramos a forma de $\psi(\nu_2)$ em (1.1.8) substituindo ν_2 por (1.1.20), tal que

$$\nu_1 = -\frac{\partial c(\lambda_1, \nu_2)}{\partial \lambda_1} = \frac{1}{\nu_2} - \frac{1}{2\lambda_1}. \quad (1.1.21)$$

Assim, $\psi(\nu_2) = 1/\nu_2$, implicando que $\psi'(\nu_2) = -1/\nu_2^2$, e $G'(\lambda_1) = -1/2\lambda_1$.

Finalmente, temos que λ_2 pode ser escrito na forma da Suposição 1.1.1 como:

$$\lambda_2^* = -\lambda_1\psi'(\nu_2) = \frac{\lambda_1}{\nu_2^2}. \quad (1.1.22)$$

A partir de (1.1.9), temos que

$$Var[U_2(W)] = Var[W] = \frac{\nu_2^3}{-2\lambda_1}, \quad (1.1.23)$$

com $Var[W] > 0$, pois $\lambda_1 = -a/2 < 0$ uma vez que $a > 0$.

Logo, se tomarmos $\nu_2 = \mu$ e $\lambda_1 = -\phi/2$ temos que $(\lambda_1, \lambda_2^*) = (-\phi/2, -\phi/2\mu^2)$ e reescrevemos a função densidade em (1.1.19) em função de μ e ϕ como

$$\begin{aligned} f_W(w|(\mu, \phi)) &= (2\pi)^{-1/2} w^{-3/2} \exp \left\{ -\frac{\phi}{2\mu^2} w - \frac{\phi}{2w} + \left(\frac{\phi^2}{\mu^2} \right)^{1/2} + \frac{1}{2} \log(\phi) \right\} \\ &= \sqrt{\frac{\phi}{2\pi w^3}} \exp \left\{ -\phi \frac{(w - \mu)^2}{2\mu^2 w} \right\}, \quad w > 0, \mu > 0, \phi > 0, \end{aligned} \quad (1.1.24)$$

a qual satisfaz (1.1.7) com $h(w) = w^{-3/2}$, $\eta_1(\mu, \phi) = -\phi/2$, $\eta_2(\mu, \phi) = -\phi/2\mu^2$, $U_1(w) = 1/w$, $U_2(w) = w$ e $b(\mu, \phi) = \phi/\mu + 1/2 \log(\phi)$, tal que

$$E[W] = \mu \quad \text{e} \quad Var[W] = \frac{\mu^3}{\phi}. \quad (1.1.25)$$

A Figura 1.2 apresenta a função densidade dada em (1.1.24) para alguns valores de μ e ϕ . Em (a), notamos que ϕ controla a assimetria da distribuição assim como sua variância, que diminui quando ϕ aumenta. Em (b), notamos que à medida que μ aumenta, a variância também aumenta.

(iii) Distribuição Lognormal

Neste caso, utilizaremos o fato de que as Suposições 1.1.1 e 1.1.2 valem para a função densidade Normal e para transformações 1 a 1 a partir dela, como mostrado em Bar-Lev e Reiser

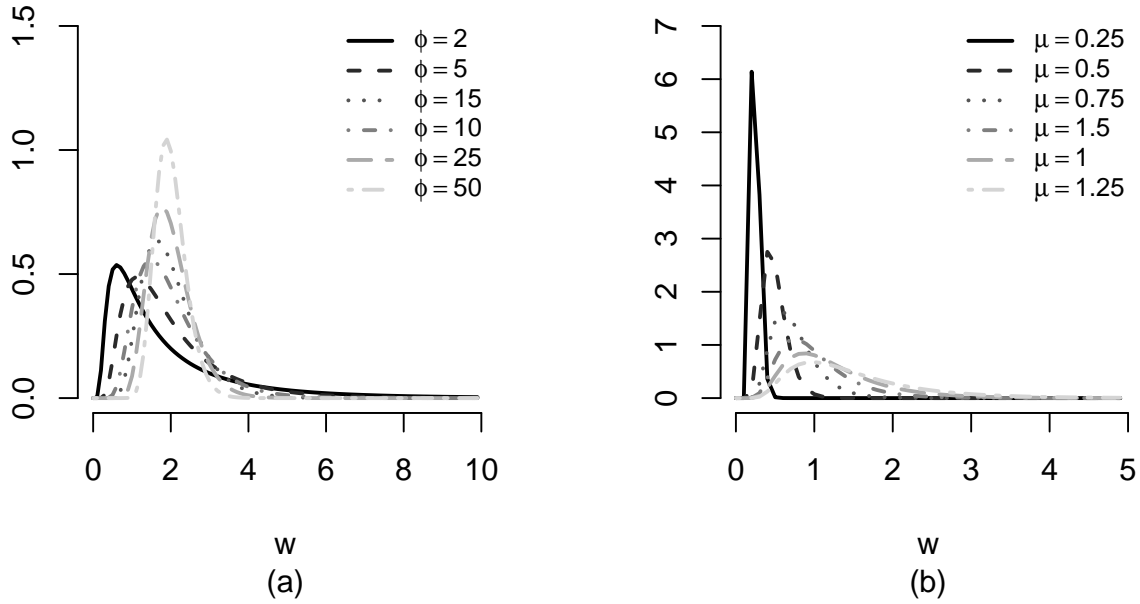


Figura 1.2: Função densidade $IG(\mu, \phi)$ com $\mu = 2$ em (a) e $\phi = 5$ em (b).

(1982).

Seja Z uma variável aleatória com densidade Normal com média a e variância b^2 . Então $W = \exp(Z)$ tem distribuição Lognormal com parâmetros a e b^2 , ou seja, $W \sim \text{LN}(a, b^2)$ e sua função densidade é dada por

$$f_W(w|(a, b^2)) = \frac{1}{b w \sqrt{2\pi}} \exp \left\{ -\frac{(\log(w) - a)^2}{2b^2} \right\}, \quad w > 0, \quad a \in \mathbb{R}, \quad b > 0. \quad (1.1.26)$$

Representando (1.1.26) na forma (1.1.4), temos

$$f_W(w|\boldsymbol{\lambda}) = \frac{1}{\sqrt{2\pi}} w^{-1} \exp \left\{ \lambda_1 (\log(w))^2 + \lambda_2 \log(w) + \frac{\lambda_2^2}{4\lambda_1^2} + \frac{1}{2} \log(-2\lambda_1) \right\}, \quad w > 0, \quad (1.1.27)$$

com $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ com $\lambda_1 = -1/2b^2$ e $\lambda_2 = a/b^2$, $h(w) = w^{-1}$, $U_1(w) = (\log(w))^2$, $U_2(w) = \log(w)$ e $c(\lambda_1, \lambda_2) = \lambda_2^2/4\lambda_1 + 1/2 \log(-2\lambda_1)$.

Como $U_2(W) = \log(W)$, então $\nu_2 = E[U_2(W)] = E[\log(W)]$. Logo, não podemos usar

diretamente a Suposição 1.1.1.

A função densidade de Z pode ser escrita como

$$f_Z(z|(a, b^2)) = \frac{1}{b\sqrt{2\pi}} \exp\left\{-\frac{(z-a)^2}{2b^2}\right\}, \quad z \in \mathbb{R}, \quad a \in \mathbb{R}, \quad b > 0, \quad (1.1.28)$$

e satisfaz (1.1.7) com $(\mu, \phi) = (a, b^2)$ e $h(z) = 1$, $\eta_1(\mu, \phi) = -1/2\phi$, $\eta_2(\mu, \phi) = \mu/\phi$, $U_1(z) = z^2$, $U_2(z) = z$ e $b(\mu, \phi) = (\mu/\phi)^2/4(-1/2\phi) + (1/2) \log(1/\phi)$, ou seja, temos que a distribuição é parametrizada diretamente pela média $\mu = a = E[W]$ e pela variância $\phi = b^2 = Var[W]$, pertencendo a família de distribuições em (1.1.7).

Além disso, $E[Z] = a$ e $Var[Z] = b^2$, então sua função geradora de momentos é dada por

$$M_Z(t) = E[\exp(tZ)] = \exp\left\{at + \frac{t^2 b^2}{2}\right\}, \quad t \in \mathbb{R}. \quad (1.1.29)$$

Assim, se Z tem distribuição normal e $W = \exp(Z)$, temos que a esperança e a variância de W , expressas através da esperança e variância de Z , são dadas por

$$E[W] = E[\exp(Z)] = M_Z(1) = \exp\left(E[Z] + \frac{1}{2}Var[Z]\right) \quad \text{e} \quad (1.1.30a)$$

$$\begin{aligned} Var[W] &= Var[\exp(Z)] = E[\exp^2(Z)] - E^2[\exp(Z)] = M_Z(2) - (M_Z(1))^2 \\ &= \exp(2E[Z] + 2Var[Z]) - (\exp(E[Z] + (1/2)Var[Z]))^2 \\ &= \exp(2E[Z] + Var[Z]) \exp(Var[Z]) - \exp(2E[Z] + Var[Z]) \\ &= \exp(2E[Z] + Var[Z])(\exp(Var[Z]) - 1), \end{aligned} \quad (1.1.30b)$$

respectivamente.

Resolvendo o sistema formado pelas equações (1.1.30a) e (1.1.30b), temos que a esperança e variância de Z podem ser escritas em função da esperança e variância de W , ou seja,

$$E[Z] = \log(E[W]) - \frac{1}{2} \log\left(1 + \frac{Var[W]}{E^2[W]}\right) \quad \text{e} \quad Var[Z] = \log\left(1 + \frac{Var[W]}{E^2[W]}\right). \quad (1.1.31)$$

Definindo

$$\mu_W = \exp\left(E[Z] + \frac{1}{2}Var[Z]\right) \quad \text{e} \quad \phi_W = (\exp(Var[Z]) - 1)^{-1}, \quad (1.1.32)$$

então por (1.1.30a) e (1.1.30b), temos

$$E[W] = \mu_W \quad \text{e} \quad Var[W] = \frac{\mu_W^2}{\phi_W}. \quad (1.1.33)$$

Assim,

$$E[Z] = \log(\mu_W) - \frac{1}{2} \log\left(\frac{1}{\phi_W} + 1\right) \quad \text{e} \quad Var[Z] = \log\left(\frac{1}{\phi_W} + 1\right). \quad (1.1.34)$$

Agora, se $Z \sim N(a^*, b^{2*})$ e $W = \exp(Z)$, pelo método do Jacobiano com $a^* = E(Z) = \log(\mu_W) - 1/2 \log(1/\phi_W + 1)$ e $b^{2*} = Var(Z) = \log(1/\phi_W + 1)$, temos que

$$\begin{aligned} f_W(w | (\mu_W, \phi_W)) &= f_Z(\log(w) | (a^*, b^{2*})) \times \frac{1}{w} \\ &= \frac{1}{w \sqrt{2\pi \log(\phi_W^{-1} + 1)}} \exp\left\{-\frac{(\log(w) - \log(\mu_W) + (1/2) \log(\phi_W^{-1} + 1))^2}{2 \log(\phi_W^{-1} + 1)}\right\}, \end{aligned} \quad (1.1.35)$$

que é a função densidade Lognormal, com forma em (1.1.26) com $a = a^*$ e $b = b^*$, de parâmetros μ_W, ϕ_W tal que $\mu_W \in \mathbb{R}$, $\phi_W > 0$. Finalmente,

$$E[W] = \mu_W \quad \text{e} \quad Var[W] = \frac{\mu_W^2}{\phi_W}. \quad (1.1.36)$$

A densidade em (1.1.35) satisfaz (1.1.7), com $h(w) = w^{-1}$, $\eta_1(\mu_W, \phi_W) = -1/2 \log(\phi_W^{-1} + 1)$, $\eta_2(\mu_W, \phi_W) = (\log(\mu_W) - (1/2) \log(\phi_W^{-1} + 1))/\log(\phi_W^{-1} + 1)$, $U_1(w) = (\log(w))^2$, $U_2(w) = \log(w)$ e $b(\mu_W, \phi_W) = -(\log(\mu_W) - (1/2) \log(\phi_W^{-1} + 1))^2/2 \log(\phi_W^{-1} + 1) - (1/2) \log(\log(\phi_W^{-1} + 1))$.

A Figura 1.3 apresenta a função densidade de $W \sim LN(\mu, \phi)$ dada por (1.1.35) para alguns valores de μ e ϕ . Em (a), podemos notar que ϕ controla a assimetria da distribuição assim como sua variância, que diminui quando ϕ aumenta. Em (b), notamos que à medida que μ

aumenta, a variância da distribuição aumenta.

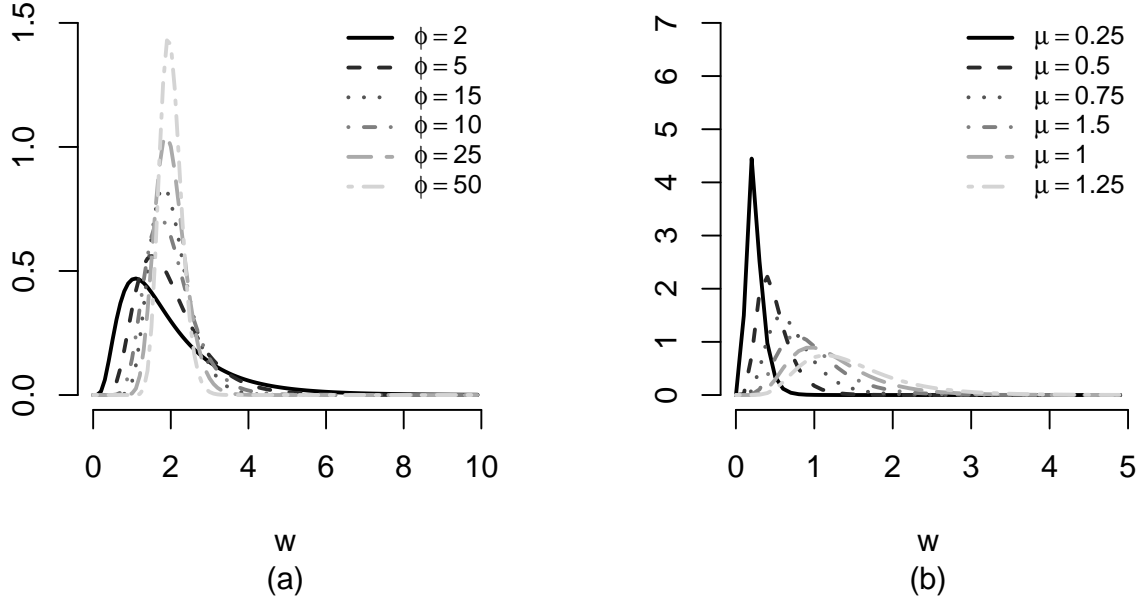


Figura 1.3: Função densidade $LN(\mu, \phi)$ com $\mu = 2$ em (a) e $\phi = 5$ em (b).

1.1.3 Modelo de regressão para respostas positivas aumentadas em zero com efeitos aleatórios

Sejam $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ vetores aleatórios independentes tal que $\mathbf{y}_i = (y_{i1}, \dots, y_{in_i})$ é a realização do vetor aleatório \mathbf{Y}_i e representa as respostas da i -ésima unidade amostral, tal que $y_{ij} \in [0, \infty)$, para $j = 1, \dots, n_i$.

Seja $\mathbf{U}_i = (\mathbf{U}_{1,i}^\top, \mathbf{U}_{2,i}^\top)$ um vetor q -dimensional de efeitos aleatórios, onde $\mathbf{U}_{1,i}$ e $\mathbf{U}_{2,i}$ são vetores de dimensão $q_\mu \times 1$ e $q_\phi \times 1$, respectivamente.

Condicional em $\mathbf{U}_i = \mathbf{u}_i$, $Y_{ij}|\mathbf{u}_i$ tem distribuição de probabilidade $g(y_{ij}|\boldsymbol{\lambda}_{ij}, p_{ij})$ na forma (1.1.1), com $f(\cdot|\boldsymbol{\lambda})$ pertencente à família exponencial descrita em (1.1.7) com $\boldsymbol{\lambda}_{ij} = (\mu_{ij}, \phi)$. Aqui, $\mu_{ij} = E[Y_{ij}|Y_{ij} > 0, \mathbf{U}_i = \mathbf{u}_i]$ e ϕ é o parâmetro de precisão. Nesse estudo, ϕ é considerado homogêneo.

Modelamos μ_{ij} utilizando a função logaritmo, ou seja,

$$\log(\mu_{ij}) = \mathbf{x}_{ij,\mu} \boldsymbol{\beta}_\mu + \mathbf{z}_{ij,\mu} \mathbf{u}_{1,i}. \quad (1.1.37)$$

Adicionalmente, modelamos $p_{ij} = P(Y_{ij} = 0 | \mathbf{U}_i = \mathbf{u}_i)$ utilizando a função logito, isto é, $\text{logito}(x) = \log(x/(1-x))$. Em particular, consideramos

$$\text{logito}(p_{ij}) = \mathbf{x}_{ij,p} \boldsymbol{\beta}_p + \mathbf{z}_{ij,p} \mathbf{u}_{2,i}, \quad (1.1.38)$$

onde $\boldsymbol{\mu}_i = (\mu_{i1}, \dots, \mu_{in_i})^\top$ e $\mathbf{p}_i = (p_{i1}, \dots, p_{in_i})^\top$. Aqui, $\mathbf{x}_{ij,\mu}$ e $\mathbf{x}_{ij,p}$ representam a ij -ésima linha das matrizes de desenho \mathbf{X}_μ e \mathbf{X}_p , respectivamente, as quais são assumidas ter posto completo e dimensões $J \times r_\mu$ e $J \times r_p$, com $J = \sum_{i=1}^n n_i$, e estão associadas aos vetores de parâmetros $\boldsymbol{\beta}_\mu$ e $\boldsymbol{\beta}_p$ de dimensões $r_\mu \times 1$ e $r_p \times 1$, respectivamente. Além disso, $\mathbf{z}_{ij,\mu}$ e $\mathbf{z}_{ij,p}$ representam a ij -ésima linha das matrizes de desenho \mathbf{Z}_μ e \mathbf{Z}_p , respectivamente, as quais são assumidas ter posto completo e dimensões $J \times q_\mu$ e $J \times q_p$ e estão associadas aos vetores de efeitos aleatórios $\mathbf{u}_{1,i}$ e $\mathbf{u}_{2,i}$ de dimensões $q_\mu \times 1$ e $q_p \times 1$, respectivamente.

Assumimos que $\mathbf{U}_i = (\mathbf{U}_{1,i}^\top, \mathbf{U}_{2,i}^\top)^\top \stackrel{i.i.d}{\sim} N_{q_\mu+q_p}(\mathbf{0}, \boldsymbol{\Sigma})$, onde $\boldsymbol{\Sigma}$ é a matriz de covariâncias de \mathbf{U}_i , os quais são mutuamente independentes. Permitimos aqui que exista uma correlação entre o vetor de efeitos aleatórios $\mathbf{U}_{2,i}$, associado à $\text{logito}(p_{ij})$, e o vetor de efeitos aleatórios $\mathbf{U}_{1,i}$, associado à $\log(\mu_{ij})$. Intuitivamente, esperamos que a correlação seja negativa, já que uma média alta implica que a probabilidade do zero deve ser baixa, e vice-versa (Olsen e Schafer, 2001).

Sob este modelo, analogamente à (1.1.2), temos que a média e a variância de Y_{ij} condicionadas nos efeitos aleatórios $\mathbf{U}_i = \mathbf{u}_i$ são dadas por

$$E(Y_{ij} | \mathbf{U}_i = \mathbf{u}_i) = (1-p_{ij}) \mu_{ij} \quad \text{e} \quad \text{Var}(Y_{ij} | \mathbf{U}_i = \mathbf{u}_i) = (1-p_{ij}) \sigma^2(\mu_{ij}, \phi) + p_{ij} (1-p_{ij}) \mu_{ij}^2, \quad (1.1.39)$$

onde $\mu_{ij} = E[Y_{ij} | Y_{ij} > 0, \mathbf{U}_i = \mathbf{u}_i]$ e $\sigma^2(\mu_{ij}, \phi) = \text{Var}[Y_{ij} | Y_{ij} > 0, \mathbf{U}_i = \mathbf{u}_i]$.

Em estudos sobre consumo habitual, um dos interesses é estimar a variabilidade intra indivíduo e também a heterogeneidade entre o consumo médio dos indivíduos (Verly Junior et al., 2010). Buscando estimar estas duas fontes de variabilidade, propomos uma extensão do modelo proposto em Rodrigues-Motta et al. (2015), acrescentando à $\log(\mu_{ij})$ efeitos aleatórios que permitam estimar a variabilidade intra e entre indivíduos. Assim, a partir de (1.1.37) e

(1.1.38), modelamos a média μ_{ij} e a probabilidade de valor zero p_{ij} através de

$$\begin{aligned} \log(\mu_{ij}) &= \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \\ \text{logito}(p_{ij}) &= \mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p + u_{2,i} \end{aligned} \quad , \quad (1.1.40)$$

onde o vetor $\mathbf{u}_{ij} = (u_{1,i}, b_{ij}, u_{2,i})$ é uma realização do vetor $\mathbf{U}_{ij} = (U_{1,i}, B_{ij}, U_{2,i})$, tal que $B_{ij} \stackrel{i.i.d}{\sim} N(0, \sigma_b^2)$ independente de $(U_{1,i}, U_{2,i})^\top \stackrel{i.i.d}{\sim} N_2(\mathbf{0}, \boldsymbol{\Sigma}_u)$, onde

$$\boldsymbol{\Sigma}_u = \begin{pmatrix} \sigma_{u_1}^2 & \rho\sigma_{u_1}\sigma_{u_2} \\ \rho\sigma_{u_1}\sigma_{u_2} & \sigma_{u_2}^2 \end{pmatrix}. \quad (1.1.41)$$

Os efeitos aleatórios $U_{1,i}$ e $U_{2,i}$ acomodam a heterogeneidade entre $\log(\mu_{ij})$ e $\text{logito}(p_{ij})$, respectivamente, onde $\sigma_{u_1}^2$ e $\sigma_{u_2}^2$ são as variâncias de $U_{1,i}$ e $U_{2,i}$, respectivamente. Aqui, ρ acomoda a correlação entre $U_{1,i}$ e $U_{2,i}$, induzindo a uma correlação entre $\log(\mu_{ij})$ e $\text{logito}(p_{ij})$.

Finalmente, note que (1.1.40) é um caso particular das equações em (1.1.37) e (1.1.38), onde $\mathbf{U}_i = (U_{1,i}, B_{i1}, \dots, B_{in_i}, U_{2,i})^\top$ tal que $\mathbf{U}_i \stackrel{i.i.d}{\sim} N_{n_i+2}(\mathbf{0}, \boldsymbol{\Sigma})$, onde $\boldsymbol{\Sigma}$ é dada por

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{u_1}^2 & \mathbf{0} & \rho\sigma_{u_1}\sigma_{u_2} \\ \mathbf{0} & \sigma_b^2 \mathbf{I}_{n_i} & \mathbf{0} \\ \rho\sigma_{u_1}\sigma_{u_2} & \mathbf{0} & \sigma_{u_2}^2 \end{pmatrix} \quad , \quad (1.1.42)$$

onde \mathbf{I}_{n_i} é uma matriz identidade de ordem n_i .

Para facilitar o processo de estimação, rearranjamos $\mathbf{U}_i = (U_{1,i}, U_{2,i}, B_{i1}, \dots, B_{in_i})^\top$ tal que $\mathbf{U}_i \stackrel{i.i.d}{\sim} N_{n_i+2}(\mathbf{0}, \boldsymbol{\Sigma})$, onde agora $\boldsymbol{\Sigma}$ é dada por

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{u_1}^2 & \rho\sigma_{u_1}\sigma_{u_2} & \mathbf{0} \\ \rho\sigma_{u_1}\sigma_{u_2} & \sigma_{u_2}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_b^2 \mathbf{I}_{n_i} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma}_u & \mathbf{0} \\ \mathbf{0} & \sigma_b^2 \mathbf{I}_{n_i} \end{pmatrix}, \quad (1.1.43)$$

para $\boldsymbol{\Sigma}_u = \begin{pmatrix} \sigma_{u_1}^2 & \rho\sigma_{u_1}\sigma_{u_2} \\ \rho\sigma_{u_1}\sigma_{u_2} & \sigma_{u_2}^2 \end{pmatrix}.$

Apresentamos a seguir a função de verossimilhança para o modelo (1.1.3), considerando

(1.1.40) e (1.1.43).

1.1.4 Verossimilhança para respostas positivas aumentadas em zero

Seja $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ vetores aleatórios independentes, com $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{in_i})$. Além disso, assumamos que, condicionado em $\mathbf{U}_i = \mathbf{u}_i$, Y_{i1}, \dots, Y_{in_i} são independentes. Assim, a verossimilhança de $\boldsymbol{\theta} = (\boldsymbol{\beta}_\mu, \boldsymbol{\beta}_p, \phi, \boldsymbol{\Sigma})$ pode ser expressa como

$$L(\boldsymbol{\theta}|\mathbf{y}, \mathbf{X}_\mu, \mathbf{X}_p) = \prod_{i=1}^n \int \prod_{j=1}^{n_i} g(y_{ij}|\mu_{ij}, \phi, p_{ij}) \phi_{\mathbf{U}_i}(\mathbf{u}_i|\boldsymbol{\Sigma}) d\mathbf{u}_i, \quad (1.1.44)$$

onde $g(y_{ij}|\mu_{ij}, \phi, p_{ij})$ é da forma (1.1.1) com $f(y_{ij}|\mu_{ij}, \phi)$ pertencente à família exponencial de dois parâmetros parametrizada como em (1.1.7) e $\phi_{\mathbf{U}_i}(\cdot)$ é a densidade dos efeitos aleatórios em (1.1.40). Aqui, $\mu_{ij} = \exp \left\{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \right\}$ e $p_{ij} = \exp \left\{ \mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p + u_{2,i} \right\} / \left(1 + \exp \left\{ \mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p + u_{2,i} \right\} \right)$ são funções dos efeitos fixos e aleatórios como em (1.1.40).

Assim,

$$\begin{aligned} L(\boldsymbol{\theta}|\mathbf{y}, \mathbf{X}_\mu, \mathbf{X}_p) &= \prod_{i=1}^n \int \prod_{j=1}^{n_i} \left(\frac{p_{ij}}{1 - p_{ij}} \right)^{\mathbf{I}_{\{y_{ij}=0\}}} (1 - p_{ij}) f(y_{ij}|\mu_{ij}, \phi)^{1 - \mathbf{I}_{\{y_{ij}=0\}}} \\ &\quad \times \frac{1}{2\pi|\boldsymbol{\Sigma}_{\mathbf{u}}|^{1/2}} \exp \left\{ -\frac{1}{2} (u_{1,i} \ u_{2,i})^\top \boldsymbol{\Sigma}_{\mathbf{u}}^{-1} (u_{1,i} \ u_{2,i}) \right\} \frac{1}{(2\pi\sigma_b^2)^{1/2}} \exp \left\{ -\frac{1}{2\sigma_b^2} b_{ij}^2 \right\} d\mathbf{u}_i. \end{aligned} \quad (1.1.45)$$

A seguir, abordamos o método inferencial para a estimação do vetor de parâmetros $\boldsymbol{\theta}$.

Capítulo 2

Método Inferencial

A integral múltipla em (1.1.45) não pode ser resolvida analiticamente. No entanto, avaliar as integrais na função de verossimilhanças com o objetivo de se calcular estimativas via método de máxima verossimilhança pode apresentar problemas numéricos quando a dimensão dos efeitos aleatórios é muito grande. Além disso, quando o número de observações é pequeno, a teoria assintótica baseada no método de máxima verossimilhança pode não se aplicar.

Olsen e Schafer (2001) abordam, no contexto de modelos de duas partes para respostas semi-contínuas com efeitos aleatórios, algoritmos do tipo EM e a quase-verossimilhança penalizada, lidando com integrais de alta dimensão através da aproximação da verossimilhança por Laplace e sua maximização por um algoritmo Escore de Fisher aproximado. Além da aproximação por Laplace, a aproximação direta da verossimilhança pode ser feita utilizando quadraturas do tipo Gauss-Hermite e Gaussianas adaptativas (Givens e Hoeting, 2012). Uma revisão e discussão dessas técnicas pode ser encontrada em McCulloch Charles e Searle (2001).

Uma forma de contornar este problema é utilizar métodos de Monte Carlo via Cadeias de Markov (MCMC), como por exemplo o amostrador de Gibbs (Gelfand e Smith, 1990).

No contexto dos modelos para respostas positivas aumentadas em zero com efeitos aleatórios apresentado no capítulo anterior, temos que a dimensão dos efeitos aleatórios pode trazer certos desafios computacionais para o ajuste desses modelos. Como em (1.1.44), a forma da verossimilhança pode depender de uma integral de grande dimensão (dependendo da dimensão dos efeitos aleatórios) tornando difícil avaliar e também maximizar essa função numericamente.

Assim, neste estudo, adotamos uma abordagem Bayesiana para a estimação dos parâmetros e predição dos efeitos aleatórios, motivados pela simplicidade computacional.

2.1 Inferência Bayesiana

A inferência Bayesiana considera, diferentemente da inferência frequentista, os parâmetros desconhecidos como variáveis aleatórias. Nessa metodologia, utiliza-se os dados observados para atualizar o conhecimento inicial (a priori) sobre os parâmetros, definido através de funções de distribuição de probabilidade para os parâmetros chamadas de distribuições a priori.

Essas distribuições devem expressar o conhecimento inicial ou a incerteza sobre o parâmetro antes que os dados sejam observados. Em geral, não temos informações a priori disponíveis, levando à necessidade de se especificar distribuições que não afetem o conhecimento atualizado pelos dados (a posteriori), permitindo que a informação trazida pelos dados através da verossimilhança seja dominante. Esse tipo de distribuição é chamada de priori vaga ou não informativa (Ntzoufras, 2011).

Assim, o interesse é poder amostrar da distribuição a posteriori dos parâmetros, considerando a informação a partir dos dados observados e da distribuição dos parâmetros a priori.

Seja (Y_1, \dots, Y_n) é uma amostra aleatória com função densidade conjunta dada por $f(\mathbf{y}|\boldsymbol{\theta})$ e seja $\boldsymbol{\theta}$ o vetor de parâmetros desconhecidos com distribuição a priori dada por $\pi(\boldsymbol{\theta})$. Pelo Teorema de Bayes, a distribuição a posteriori de $\boldsymbol{\theta}$, $\pi(\boldsymbol{\theta}|\mathbf{y})$, dado que a amostra \mathbf{y} foi observada, é dada por

$$\pi(\boldsymbol{\theta}|\mathbf{y}) = \frac{f(\mathbf{y}, \boldsymbol{\theta})}{f(\mathbf{y})} = \frac{\pi(\boldsymbol{\theta})f(\mathbf{y}|\boldsymbol{\theta})}{\int f(\mathbf{y}, \boldsymbol{\theta})d\boldsymbol{\theta}} = \frac{\pi(\boldsymbol{\theta})f(\mathbf{y}|\boldsymbol{\theta})}{\int f(\mathbf{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}}. \quad (2.1.1)$$

Uma vez que o denominador em (2.1.1) não depende de $\boldsymbol{\theta}$, temos que, a não ser por uma constante de normalidade, $\pi(\boldsymbol{\theta}|\mathbf{y})$ pode ser escrita como

$$\pi(\boldsymbol{\theta}|\mathbf{y}) \propto \pi(\boldsymbol{\theta})f(\mathbf{y}|\boldsymbol{\theta}). \quad (2.1.2)$$

Representando a posteriori a partir da verossimilhança, temos que

$$\pi(\boldsymbol{\theta}|\mathbf{y}) \propto \pi(\boldsymbol{\theta})L(\boldsymbol{\theta}|\mathbf{y}). \quad (2.1.3)$$

Sem perda de generalidade, seja $\boldsymbol{\theta} = \theta$. No contexto de estimação dos parâmetros, seja $\delta(\mathbf{Y})$ um estimador para θ com distribuição a posteriori dada por $\pi(\theta|\mathbf{y})$. Seja $l(\theta, a)$ a função de perda associada à estimativa a de θ quando o verdadeiro valor do parâmetro é θ , de modo que quanto maior a distância entre a e θ , maior é a perda. Para cada estimativa a , a perda esperada a posteriori é dada por $E[l(\theta, a)|\mathbf{y}] = \int l(\theta, a)\pi(\theta|\mathbf{y})d\theta$. Seja $\delta^*(\mathbf{y})$ o valor de a para o qual a perda esperada é mínima para todos os valores possíveis de \mathbf{y} . A função $\delta^*(\mathbf{Y})$ é chamada de

estimador de Bayes para θ e $\delta^*(\mathbf{y})$ o valor do estimador de Bayes tal que, para todos os valores possíveis de \mathbf{y}

$$E[l(\theta, \delta^*(\mathbf{y}))|\mathbf{y}] = \min_a E[l(\theta, a)|\mathbf{y}].$$

Diferentes funções de perda podem ser consideradas. Por exemplo, se consideramos a função de perda quadrática $l(\theta, a) = (\theta - a)^2$, então o estimador de Bayes é dado pela média a posteriori $E[\theta|\mathbf{y}]$. Se utilizarmos a função de perda absoluta $l(\theta, a) = |\theta - a|$, então o estimador de Bayes é dado pela mediana a posteriori. Se considerarmos funções que associam uma perda fixa como a função de perda 0 – 1 dada por

$$l(\theta, a) = \begin{cases} 1, & \text{se } |\theta - a| > \epsilon \\ 0, & \text{se } |\theta - a| \leq \epsilon \end{cases},$$

para um certo $\epsilon > 0$, então o estimador de Bayes é a moda a posteriori. Neste estudo, consideraremos a função de perda quadrática, e assim a média a posteriori será considerada como estimador de um parâmetro de interesse.

Em geral, não é simples obter amostras da distribuição a posteriori quando sua forma não é conhecida. Para aproximarmos essas distribuições podemos usar métodos de simulação do tipo Monte Carlo via Cadeias de Markov (MCMC), os quais são baseados em cadeias de Markov (Givens e Hoeting, 2012), garantindo que após um período de aquecimento, a distribuição dos pontos da cadeia tem aproximadamente a distribuição a posteriori de interesse. Em geral, após o período de aquecimento da cadeia, as amostras são obtidas considerando-se um certo espaçamento, de forma a diminuir a correlação entre as observações.

Uma forma de construirmos uma cadeia de Markov com distribuição limite (2.1.1) é através de um processo de simulação iterativo da distribuição de um bloco de parâmetro, condicionado nos dados e no bloco de parâmetros restantes, garantindo que amostras de todas as componentes de $\boldsymbol{\theta}$ tenham sido obtidas. A esta distribuição condicional dá-se o nome de distribuição condicional completa.

Quando no processo iterativo não for possível amostrar diretamente da distribuição condicional, utilizamos técnicas como o algoritmo de Metropolis-Hastings (MH) (Givens e Hoeting, 2012). Uma vantagem deste método é que não é necessário conhecer a constante normalizadora da distribuição, que na maioria das vezes é muito difícil de ser obtida.

Neste estudo, para simular da distribuição a posteriori dos parâmetros e dos efeitos aleatórios, usaremos métodos de MCMC via amostrador de Gibbs proposto em Geman e Geman (1984) e aplicado mais genericamente em Gelfand e Smith (1990).

2.1.1 O amostrador de Gibbs

Abordando o algoritmo de Gibbs brevemente, suponha que $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)$ e seja $\boldsymbol{\theta}_{(i)}$ o bloco de parâmetros sem θ_i , tal que $\boldsymbol{\theta}_{(i)} = (\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_p)$, $i = 1, \dots, p$. Suponha que podemos gerar da distribuição condicional completa $\pi(\theta_i | \boldsymbol{\theta}_{(i)}, \mathbf{y})$, para todo i . Um esquema de Gibbs pode ser construído da seguinte forma:

1. Definimos valores iniciais $\theta_i^{(0)}$, $i = 1, \dots, p$, e tomamos $t = 0$.
2. Amostramos $\boldsymbol{\theta}^{(t+1)}$, amostrando

$$\begin{aligned} &\theta_1^{(t+1)} \text{ de } \pi(\theta_1 | \theta_2^{(t)}, \dots, \theta_p^{(t)}, \mathbf{y}), \\ &\theta_2^{(t+1)} \text{ de } \pi(\theta_2 | \theta_1^{(t+1)}, \theta_3^{(t)}, \dots, \theta_p^{(t)}, \mathbf{y}), \\ &\vdots \\ &\theta_{p-1}^{(t+1)} \text{ de } \pi(\theta_{p-1} | \theta_1^{(t+1)}, \dots, \theta_{p-2}^{(t+1)}, \theta_p^{(t)}, \mathbf{y}), \\ &\theta_p^{(t+1)} \text{ de } \pi(\theta_p | \theta_1^{(t+1)}, \dots, \theta_{p-1}^{(t+1)}, \mathbf{y}). \end{aligned}$$

3. Incrementamos t e voltamos ao passo 2.

Depois de M iterações do processo, obtemos $(\theta_1^{(M)}, \dots, \theta_p^{(M)})$. Geman e Geman (1984) mostram que para M suficientemente grande, a distribuição conjunta de $(\theta_1^{(M)}, \dots, \theta_p^{(M)})$ (condicional em \mathbf{y}) converge em distribuição para $\pi(\theta_1, \dots, \theta_p | \mathbf{y})$. Também, para cada elemento θ_i , a distribuição marginal de $\theta_i^{(M)}$ converge em distribuição para a distribuição marginal a posteriori $\pi(\theta_i | \mathbf{y})$.

Para cada θ_i em $\boldsymbol{\theta}$, dada uma amostra de tamanho l , $(\theta_i^{(M)}, \dots, \theta_i^{(M+l)})$, $l > M$, a estimativa da esperança a posteriori por Monte Carlo é dada por $\widehat{E[\theta_i | \mathbf{y}]} = (1/l) \sum_{j=M}^{M+l} \theta_i^{(j)}$ a qual corresponde à média dos valores amostrados. Sem perda de generalidade, para qualquer função $G(\theta_i)$, $\widehat{E[G(\theta_i) | \mathbf{y}]} = (1/l) \sum_{j=M}^{M+l} G(\theta_i^{(j)})$.

Em geral, as distribuições condicionais completas não têm forma conhecida. Assim, para gerar valores dessas distribuições podemos utilizar, dentro do algoritmo de Gibbs, o algoritmo

de Metropolis-Hastings (M-H).

Querendo amostrar de $\pi(\theta_i|\boldsymbol{\theta}_{(i)}, \mathbf{y})$, o algoritmo de MH consiste dos seguintes passos:

1. Definimos o valor inicial $\theta_i^{(0)}$ e tomamos $M = 0$.
2. Definimos $\theta = \theta_i^{(M)}$ e geramos θ' da distribuição candidata $q(\theta'|\theta)$. Daí calculamos

$$\alpha = \min \left(1, \frac{\pi(\theta'|\mathbf{y})q(\theta|\theta')}{\pi(\theta|\mathbf{y})q(\theta'|\theta)} \right) = \min \left(1, \frac{f(\mathbf{y}|\theta')\pi(\theta')q(\theta|\theta')}{f(\mathbf{y}|\theta)\pi(\theta)q(\theta'|\theta)} \right)$$

e aceitamos $\theta^{(M+1)} = \theta'$ com probabilidade α , caso contrário, $\theta^{(M+1)} = \theta$.

3. Incrementamos M e voltamos ao passo 2.

Mais detalhes e propriedades do algoritmo MH podem ser encontrados em Givens e Hoeting (2012).

Para algoritmos MCMC é importante verificar a convergência das cadeias. Se a convergência for verificada, podemos considerar que a amostra obtida vêm da distribuição alvo. Algumas técnicas para acessar a convergência são discutidas em Cowles e Carlin (1996). Neste estudo, utilizamos as estatísticas \hat{R} e gráficos do tipo trace plots para as cadeias.

2.2 Análise Bayesiana do modelo de regressão para respostas positivas aumentas de zero

Seja $\boldsymbol{\theta} = (\boldsymbol{\beta}_\mu, \boldsymbol{\beta}_p, \phi, \boldsymbol{\Sigma})$ o vetor de parâmetros e \mathbf{H} o vetor de hiperparâmetros conhecidos. Por simplicidade, assumimos que os elementos em $\boldsymbol{\theta}$ são independentes a priori, tal que a distribuição a priori conjunta é dada por

$$\pi(\boldsymbol{\theta}|\mathbf{H}) = \pi(\boldsymbol{\beta}_\mu)\pi(\boldsymbol{\beta}_p)\pi(\boldsymbol{\Sigma})\pi(\phi). \quad (2.2.1)$$

Combinando a função de verossimilhança em (1.1.45) com a distribuição conjunta a priori em (2.2.1), a distribuição a posteriori dos parâmetros aumentada pelos efeitos aleatórios $\mathbf{u} =$

$(\mathbf{u}_1, \dots, \mathbf{u}_n)$, com $\mathbf{u}_i = (u_{1,i}, u_{2,i}, b_{i1}, \dots, b_{in_i})$, é definida por

$$\pi(\boldsymbol{\theta}, \mathbf{u} | \mathbf{y}, \mathbf{H}) \propto L(\boldsymbol{\theta} | \mathbf{u}, \mathbf{y}, \mathbf{X}_\mu, \mathbf{X}_p) \pi(\boldsymbol{\theta} | \mathbf{H}), \quad (2.2.2)$$

onde $L(\boldsymbol{\theta} | \mathbf{u}, \mathbf{y}, \mathbf{X}_\mu, \mathbf{X}_p)$ depende da distribuição da família exponencial em (1.1.7) parametrizada por $\mu_{ij} = \exp \left\{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \right\}$ e $p_{ij} = \exp \left\{ \mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p + u_{2,i} \right\} / \left(1 + \exp \left\{ \mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p + u_{2,i} \right\} \right)$.

Para os parâmetros de efeitos fixos do modelo $\boldsymbol{\beta}_\mu, \boldsymbol{\beta}_p$, especificamos prioris normal multivariadas, ou seja, supomos $\boldsymbol{\beta}_\mu \sim N_{q_\mu}(\mathbf{0}, \sigma_{\beta_{\mu 0}}^2 \mathbf{I}_{q_\mu})$ e $\boldsymbol{\beta}_p \sim N_{q_p}(\mathbf{0}, \sigma_{\beta_{p 0}}^2 \mathbf{I}_{q_p})$, onde $\sigma_{\beta_{\mu 0}}^2$ e $\sigma_{\beta_{p 0}}^2$ são conhecidos. Para o parâmetro de precisão ϕ , temos que $\phi \sim \text{Gama}(\delta_{\phi 0}, \gamma_{\phi 0})$, onde $\delta_{\phi 0}$ e $\gamma_{\phi 0}$ são conhecidos.

A escolha da distribuição $\pi(\boldsymbol{\Sigma} | \mathbf{H})$ vem de um estudo de prioris a partir do caso geral $(u_{1,i}, u_{2,i}) \sim N(\mathbf{0}, \boldsymbol{\Sigma}_u)$ que está descrito no Capítulo 3 (veja 3.1).

Seguindo Neelon et al. (2010) neste estudo, consideramos que

$$u_{1,i} \sim N(0, \eta_1^2) \quad \text{e} \quad u_{2,i} | u_{1,i} \sim N(\psi u_{1,i}, \eta_2^2), \quad (2.2.3)$$

tal que $(u_{1,i}, u_{2,i}) \sim N(\mathbf{0}, \boldsymbol{\Sigma}_u)$, com

$$\boldsymbol{\Sigma}_u = \begin{pmatrix} \eta_1^2 & \psi \eta_1^2 \\ \psi \eta_1^2 & \psi^2 \eta_1^2 + \eta_2^2 \end{pmatrix}, \quad (2.2.4)$$

onde os elementos da matriz $\boldsymbol{\Sigma}_u$ definida em (4.2.2) são dados por

$$\sigma_{u_1}^2 = \eta_1^2, \quad \sigma_{u_2}^2 = \psi^2 \eta_1^2 + \eta_2^2, \quad \text{e} \quad \rho = \frac{\psi \eta_1^2}{\sqrt{\eta_1^2 (\psi^2 \eta_1^2 + \eta_2^2)}}. \quad (2.2.5)$$

Assumimos que $\psi \sim N(0, \sigma_{\psi 0}^2)$ e $\eta_1^2, \eta_2^2 \sim \text{Inv-Gama}(\delta_{\eta 0}, \gamma_{\eta 0})$, com hiperparâmetros conhecidos. Além disso, consideramos $\sigma_b^2 \sim \text{Inv-Gama}(\delta_{b 0}, \gamma_{b 0})$.

Finalmente, a distribuição a priori conjunta pode ser escrita como

$$\begin{aligned}
\pi(\boldsymbol{\theta}|\mathbf{H}) &= \pi(\boldsymbol{\beta}_\mu|\sigma_{\mu 0}^2)\pi(\boldsymbol{\beta}_p|\sigma_{p 0}^2)\pi(\phi|\delta_{\phi 0}, \gamma_{\phi 0})\pi(\psi|\sigma_{\psi 0}^2)\pi(\eta_1^2|\delta_{\eta 0}, \gamma_{\eta 0})\pi(\eta_2^2|\delta_{\eta 0}, \gamma_{\eta 0})\pi(\sigma_b^2|\delta_{b 0}, \gamma_{b 0}) \\
&= \frac{1}{(2\pi)^{\frac{q_\mu}{2}}(\sigma_{\beta_\mu}^2)^{\frac{q_\mu}{2}}} \exp\left\{-\frac{1}{2\sigma_{\beta_\mu}^2}\boldsymbol{\beta}'_\mu\boldsymbol{\beta}_\mu\right\} \frac{1}{(2\pi)^{\frac{q_p}{2}}(\sigma_{\beta_p}^2)^{\frac{q_p}{2}}} \exp\left\{-\frac{1}{2\sigma_{\beta_p}^2}\boldsymbol{\beta}'_p\boldsymbol{\beta}_p\right\} \\
&\times \frac{\gamma_{\phi 0}^{\delta_{\phi 0}}}{\Gamma(\delta_{\phi 0})}\phi^{\delta_{\phi 0}-1}\exp\{-\gamma_{\phi 0}\phi\} \frac{1}{(2\pi)^{1/2}(\sigma_{\psi 0}^2)^{1/2}} \exp\left\{-\frac{1}{2\sigma_{\psi 0}^2}\psi^2\right\} \\
&\times \frac{\gamma_{\eta 0}^{\delta_{\eta 0}}}{\Gamma(\delta_{\eta 0})}(\eta_1^2)^{-\delta_{\eta 0}-1}\exp\left\{-\frac{\gamma_{\eta 0}}{\eta_1^2}\right\} \frac{\gamma_{\eta 0}^{\delta_{\eta 0}}}{\Gamma(\delta_{\eta 0})}(\eta_2^2)^{-\delta_{\eta 0}-1}\exp\left\{-\frac{\gamma_{\eta 0}}{\eta_2^2}\right\} \\
&\times \frac{\gamma_{b 0}^{\delta_{b 0}}}{\Gamma(\delta_{b 0})}(\sigma_b^2)^{-\delta_{b 0}-1}\exp\left\{-\frac{\gamma_{b 0}}{\sigma_b^2}\right\}
\end{aligned} \tag{2.2.6}$$

Considerando (2.2.3), a verossimilhança do modelo em (1.1.40) é dada por

$$\begin{aligned}
L(\boldsymbol{\theta}|\mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p) &= \prod_{i=1}^n \prod_{j=1}^{n_i} g(y_{ij}|\mu_{ij}, \phi, p_{ij})\phi_{\mathbf{U}_i}(\mathbf{u}_i|\boldsymbol{\Sigma}) \\
&= \prod_{i=1}^n \prod_{j=1}^{n_i} g(y_{ij}|\mu_{ij}, \phi, p_{ij})\phi_{U_{1,i}, U_{2,i}}(u_{1,i}, u_{2,i}|\boldsymbol{\Sigma}_{\mathbf{u}})\phi_{B_{ij}}(b_{ij}|\sigma_b^2) \\
&= \prod_{i=1}^n \prod_{j=1}^{n_i} g(y_{ij}|\mu_{ij}, \phi, p_{ij})\phi_{U_{2,i}|U_{1,i}=u_{1,i}}(u_{2,i}|u_{1,i}, \boldsymbol{\Sigma}_{\mathbf{u}})\phi_{U_{1,i}}(u_{1,i}|\eta_1^2)\phi_{B_{ij}}(b_{ij}|\sigma_b^2) \\
&= \prod_{i=1}^n \prod_{j=1}^{n_i} \left(\frac{p_{ij}}{1-p_{ij}}\right)^{\mathbf{I}_{\{y_{ij}=0\}}} (1-p_{ij})f(y_{ij}|\mu_{ij}, \phi)^{1-\mathbf{I}_{\{y_{ij}=0\}}} \\
&\times \phi_{U_{2,i}|U_{1,i}=u_{1,i}}(u_{2,i}|u_{1,i}, \boldsymbol{\Sigma}_{\mathbf{u}})\phi_{U_{1,i}}(u_{1,i}|\eta_1^2)\phi_{B_{ij}}(b_{ij}|\sigma_b^2)
\end{aligned} \tag{2.2.7}$$

e depende da distribuição $g(\cdot|\cdot)$ que é parametrizada por $\mu_{ij} = \exp\{\mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij}\}$ e $p_{ij} = \exp\{\mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p + u_{2,i}\} / (1 + \exp\{\mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p + u_{2,i}\})$. Aqui, $\phi_{U_{1,i}, U_{2,i}}(\cdot)$, $\phi_{B_{ij}}(\cdot)$, $\phi_{U_{2,i}|U_{1,i}}(\cdot)$ e $\phi_{U_{1,i}}(\cdot)$ são as densidades de $(U_{1,i}, U_{2,i})$, B_{ij} , $U_{2,i}|U_{1,i} = u_{1,i}$ e $U_{1,i}$, respectivamente. Neste estudo, assumimos $\phi(\cdot)$ igual a distribuição Normal. Estamos interessados nos casos onde $f(y_{ij}|\mu_{ij}, \phi)$ é a densidade da distribuição Gama, Inversa Gaussiana e Lognormal, tal que (2.2.7) é escrita em cada um dos casos, dados a seguir.

Para a distribuição Gama, temos

$$\begin{aligned}
L(\boldsymbol{\theta}|\mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p) &= \prod_{i=1}^n \prod_{j=1}^{n_i} \left[\exp \left\{ \mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p + u_{2,i} \right\} \right]^{\mathbf{I}_{\{y_{ij}=0\}}} \frac{1}{1 + \exp(\mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p + u_{2,i})} \\
&\times \left[\frac{1}{\Gamma(\phi)} \left(\frac{\phi}{\exp \left\{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \right\}} \right)^\phi y_{ij}^{\phi-1} \exp \left\{ -\frac{\phi y_{ij}}{\exp \left\{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \right\}} \right\} \right]^{1-\mathbf{I}_{\{y_{ij}=0\}}} \\
&\times \frac{1}{(2\pi\eta_2^2)^{1/2}} \exp \left\{ -\frac{1}{2\eta_2^2} (u_{2,i} - \psi u_{1,i})^2 \right\} \frac{1}{(2\pi\eta_1^2)^{1/2}} \exp \left\{ -\frac{1}{2\eta_1^2} u_{1,i}^2 \right\} \\
&\times \frac{1}{(2\pi\sigma_b^2)^{1/2}} \exp \left\{ -\frac{1}{2\sigma_b^2} b_{ij}^2 \right\}.
\end{aligned} \tag{2.2.8}$$

Para a distribuição Inversa Gaussiana, temos

$$\begin{aligned}
L(\boldsymbol{\theta}|\mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p) &= \prod_{i=1}^n \prod_{j=1}^{n_i} \left[\exp \left\{ \mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p + u_{2,i} \right\} \right]^{\mathbf{I}_{\{y_{ij}=0\}}} \frac{1}{1 + \exp(\mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p + u_{2,i})} \\
&\times \left[\sqrt{\frac{\phi}{2\pi y_{ij}^3}} \exp \left\{ -\phi \frac{\left(y_{ij} - \exp \left\{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \right\} \right)^2}{2 \exp \left\{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \right\}^2 y_{ij}} \right\} \right]^{1-\mathbf{I}_{\{y_{ij}=0\}}} \\
&\times \frac{1}{(2\pi\eta_2^2)^{1/2}} \exp \left\{ -\frac{1}{2\eta_2^2} (u_{2,i} - \psi u_{1,i})^2 \right\} \frac{1}{(2\pi\eta_1^2)^{1/2}} \exp \left\{ -\frac{1}{2\eta_1^2} u_{1,i}^2 \right\} \\
&\times \frac{1}{(2\pi\sigma_b^2)^{1/2}} \exp \left\{ -\frac{1}{2\sigma_b^2} b_{ij}^2 \right\}.
\end{aligned} \tag{2.2.9}$$

E finalmente, para a distribuição Lognormal, temos

$$\begin{aligned}
L(\boldsymbol{\theta}|\mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p) &= \prod_{i=1}^n \prod_{j=1}^{n_i} \left[\exp \left\{ \mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p + u_{2,i} \right\} \right]^{\mathbf{I}_{\{y_{ij}=0\}}} \frac{1}{1 + \exp(\mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p + u_{2,i})} \\
&\times \left[\frac{1}{y_{ij} \sqrt{2\pi \log(\phi^{-1} + 1)}} \exp \left\{ -\frac{\left(\log \left(y_{ij} \sqrt{\phi^{-1} + 1} / \exp \left\{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \right\} \right) \right)^2}{2 \log(\phi^{-1} + 1)} \right\} \right]^{1-\mathbf{I}_{\{y_{ij}=0\}}} \\
&\times \frac{1}{(2\pi\eta_2^2)^{1/2}} \exp \left\{ -\frac{1}{2\eta_2^2} (u_{2,i} - \psi u_{1,i})^2 \right\} \frac{1}{(2\pi\eta_1^2)^{1/2}} \exp \left\{ -\frac{1}{2\eta_1^2} u_{1,i}^2 \right\} \frac{1}{(2\pi\sigma_b^2)^{1/2}} \exp \left\{ -\frac{1}{2\sigma_b^2} b_{ij}^2 \right\}.
\end{aligned} \tag{2.2.10}$$

A partir de (2.2.8), (2.2.9) e (2.2.10) e da priori conjunta dada em (2.2.6), construímos um esquema de Gibbs como descrito na Seção 2.1.1. As distribuições condicionais completas

para os parâmetros são construídas a seguir, tomando $\boldsymbol{\theta}_{(*)}$ como o conjunto de todos os outros parâmetros restantes.

A distribuição condicional completa de $\boldsymbol{\beta}_p$ não depende da forma de $f(y_{ij}|\mu_{ij}, \phi)$ e é dada por

$$\begin{aligned} \pi(\boldsymbol{\beta}_p | \mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}_{(*)}) &\propto \prod_{i=1}^n \prod_{j=1}^{n_i} \left[\exp \left\{ \mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p + u_{2,i} \right\} \right]^{\mathbf{I}_{\{y_{ij}=0\}}} \frac{1}{1 + \exp \left\{ \mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p + u_{2,i} \right\}} \\ &\times \exp \left\{ -\frac{1}{2\sigma_{\beta_p}^2} \boldsymbol{\beta}_p^\top \boldsymbol{\beta}_p \right\}, \end{aligned} \quad (2.2.11)$$

cuja forma é conhecida a não ser por uma constante de proporcionalidade. A seguir, apresentamos as demais condicionais completas de $\boldsymbol{\beta}_\mu$ para (2.2.8), (2.2.9) e (2.2.10), respectivamente.

Para a distribuição Gama, temos

$$\begin{aligned} \pi(\boldsymbol{\beta}_\mu | \mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}_{(*)}) &\propto \prod_{i=1}^n \prod_{j=1}^{n_i} \left[\left(\exp \left\{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu \right\} \right)^{-\phi} \exp \left\{ \frac{-\phi y_{ij}}{\exp \left\{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \right\}} \right\} \right]^{1 - \mathbf{I}_{\{y_{ij}=0\}}} \\ &\times \exp \left\{ -\frac{1}{2\sigma_{\beta_\mu}^2} \boldsymbol{\beta}_\mu^\top \boldsymbol{\beta}_\mu \right\}. \end{aligned} \quad (2.2.12)$$

Para a distribuição Inversa Gaussiana, temos

$$\begin{aligned} \pi(\boldsymbol{\beta}_\mu | \mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}_{(*)}) &\propto \prod_{i=1}^n \prod_{j=1}^{n_i} \left[\exp \left\{ -\phi \frac{(y_{ij} - \exp \left\{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \right\})^2}{2 \exp \left\{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \right\}^2 y_{ij}} \right\} \right]^{1 - \mathbf{I}_{\{y_{ij}=0\}}} \\ &\times \exp \left\{ -\frac{1}{2\sigma_{\beta_\mu}^2} \boldsymbol{\beta}_\mu^\top \boldsymbol{\beta}_\mu \right\}. \end{aligned} \quad (2.2.13)$$

E finalmente

$$\begin{aligned} \pi(\boldsymbol{\beta}_\mu | \mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}_{(*)}) &\propto \prod_{i=1}^n \prod_{j=1}^{n_i} \left[\exp \left\{ - \frac{\left(\log \left(y_{ij} \sqrt{\phi^{-1} + 1} / \exp \{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \} \right) \right)^2}{2 \log(\phi^{-1} + 1)} \right\} \right]^{1 - \mathbf{I}_{\{y_{ij}=0\}}} \\ &\quad \times \exp \left\{ - \frac{1}{2\sigma_{\beta_\mu}^2} \boldsymbol{\beta}_\mu^\top \boldsymbol{\beta}_\mu \right\}, \end{aligned} \quad (2.2.14)$$

para a distribuição Lognormal. A forma de (2.2.12), (2.2.13) e (2.2.14) é conhecida a não ser por uma constante de proporcionalidade.

Construiremos agora as condicionais completas para os parâmetros de dispersão. Para η_1^2 , temos

$$\begin{aligned} \pi(\eta_1^2 | \mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}_{(*)}) &\propto \prod_{i=1}^n \frac{1}{(2\pi\eta_1^2)^{1/2}} \exp \left\{ - \frac{1}{2\eta_1^2} u_{1,i}^2 \right\} (\eta_1^2)^{\delta_{\eta_0}-1} \exp \left\{ - \frac{\gamma_{\eta_0}}{\eta_1^2} \right\} \\ &= (\eta_1^2)^{-(\delta_{\eta_0}+n/2)-1} \exp \left\{ - \frac{1}{\eta_1^2} \left(\sum_{i=1}^n \frac{u_{1,i}^2}{2} + \gamma_{\eta_0} \right) \right\}, \end{aligned} \quad (2.2.15)$$

que é o núclueo de uma distribuição Inv-Gamma($\delta_{\eta_0} + n/2, \sum_{i=1}^n u_{1,i}^2/2 + \gamma_{\eta_0}$).

Para η_2^2 , temos

$$\begin{aligned} \pi(\eta_2^2 | \mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}_{(*)}) &\propto \prod_{i=1}^n \frac{1}{(2\pi\eta_2^2)^{1/2}} \exp \left\{ - \frac{1}{2\eta_2^2} (u_{2,i} - \psi u_{1,i})^2 \right\} (\eta_2^2)^{-\delta_{\eta_0}-1} \exp \left\{ - \frac{\gamma_{\eta_0}}{\eta_2^2} \right\} \\ &= (\eta_2^2)^{-(\delta_{\eta_0}+n/2)-1} \exp \left\{ - \frac{1}{\eta_2^2} \left(\sum_{i=1}^n \frac{(u_{2,i} - \psi u_{1,i})^2}{2} + \gamma_{\eta_0} \right) \right\}, \end{aligned} \quad (2.2.16)$$

que é o núclueo de uma distribuição Inv-Gamma($\delta_{\eta_0} + n/2, \sum_{i=1}^n (u_{2,i} - \psi u_{1,i})^2/2 + \gamma_{\eta_0}$).

Para ψ , temos

$$\begin{aligned} \pi(\psi|\mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}_{(*)}) &\propto \prod_{i=1}^n \exp \left\{ -\frac{1}{2\eta_2^2} (u_{2,i} - \psi u_{1,i})^2 \right\} \exp \left\{ -\frac{1}{2\sigma_{\psi 0}^2} \psi^2 \right\} \\ &= \exp \left\{ -\frac{\eta_2^2 + \sigma_{\psi 0}^2 \sum_{i=1}^n u_{1,i}^2}{2 \eta_2^2 \sigma_{\psi 0}^2} \left(\psi - \frac{2 \eta_2^2 \sigma_{\psi 0}^2 \sum_{i=1}^n u_{1,i} u_{2,i}}{\eta_2^2 + \sigma_{\psi 0}^2 \sum_{i=1}^n u_{1,i}^2} \right)^2 \right\}, \end{aligned} \quad (2.2.17)$$

que é o núclueo de uma distribuição $N(2 \eta_2^2 \sigma_{\psi 0}^2 \sum_{i=1}^n u_{1,i} u_{2,i} / (\eta_2^2 + \sigma_{\psi 0}^2 \sum_{i=1}^n u_{1,i}^2), \eta_2^2 \sigma_{\psi 0}^2 / (\eta_2^2 + \sigma_{\psi 0}^2 \sum_{i=1}^n u_{1,i}^2))$.

Para σ_b^2 , temos

$$\begin{aligned} \pi(\sigma_b^2|\mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}_{(*)}) &\propto \prod_{i=1}^n \prod_{j=1}^{n_i} \frac{1}{(2\pi\sigma_b^2)^{1/2}} \exp \left\{ -\frac{1}{2\sigma_b^2} b_{ij}^2 \right\} (\sigma_b^2)^{-\delta_{b0}-1} \exp \left\{ -\frac{\gamma_{b0}}{\sigma_b^2} \right\} \\ &= (\sigma_b^2)^{-(\delta_{b0} + \sum_{i=1}^n \sum_{j=1}^{n_i} 1/2) - 1} \exp \left\{ -\frac{1}{\sigma_b^2} \left(\sum_{i=1}^n \sum_{j=1}^{n_i} \frac{b_{ij}^2}{2} + \gamma_{b0} \right) \right\}, \end{aligned} \quad (2.2.18)$$

que é o núclueo de uma distribuição Inv-Gamma($\delta_{b0} + \sum_{i=1}^n \sum_{j=1}^{n_i} 1/2, \sum_{i=1}^n \sum_{j=1}^{n_i} b_{ij}^2/2 + \gamma_{b0}$).

A distribuição condicional do parâmetro de precisão ϕ é dada por

$$\begin{aligned} \pi(\phi|\mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}_{(*)}) &\propto \prod_{i=1}^n \prod_{j=1}^{n_i} \left[\frac{1}{\Gamma(\phi)} \left(\frac{\phi y_{ij}}{\exp \{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \}} \right)^\phi \right. \\ &\quad \times \exp \left\{ -\frac{\phi y_{ij}}{\exp \{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \}} \right\} \left. \right]^{1 - \mathbf{I}_{\{y_{ij}=0\}}} \phi^{\delta_{\phi 0}-1} \exp \{ -\gamma_{\phi 0} \phi \}, \end{aligned} \quad (2.2.19)$$

para a distribuição Gama, cuja forma é conhecida a não ser por uma constante de proporção-

nalidade, e por

$$\begin{aligned}
\pi(\phi|\mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}_{(*)}) &\propto \prod_{i=1}^n \prod_{j=1}^{n_i} \left[\sqrt{\phi} \exp \left\{ -\phi \frac{(y_{ij} - \exp \{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \})^2}{2 \exp \{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \}^2 y_{ij}} \right\} \right]^{1 - \mathbf{I}_{\{y_{ij}=0\}}} \\
&\times \phi^{\delta_{\phi 0} - 1} \exp \{ -\gamma_{\phi 0} \phi \} \\
&= \exp \left\{ -\phi \left(\sum_{i=1}^n \sum_{j=1}^{n_i} \frac{(y_{ij} - \exp \{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \})^2 \mathbf{I}_{\{y_{ij}>0\}}}{2 \exp \{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \}^2 y_{ij}} + \gamma_{\phi 0} \right) \right\} \\
&\times \phi^{\delta_{\phi 0} + \sum_{i=1}^n \sum_{j=1}^{n_i} \frac{1}{2} (1 - \mathbf{I}_{\{y_{ij}=0\}}) - 1},
\end{aligned} \tag{2.2.20}$$

para a distribuição Inversa Gaussiana, onde a condicional completa para ϕ é o núcleo de uma distribuição Gama(a, b) com

$$a = \delta_{\phi 0} + \sum_{i=1}^n \sum_{j=1}^{n_i} \frac{1}{2} (1 - \mathbf{I}_{\{y_{ij}=0\}}) \text{ e } b = \sum_{i=1}^n \sum_{j=1}^{n_i} \left(\frac{(y_{ij} - \exp \{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \})^2 \mathbf{I}_{\{y_{ij}>0\}}}{2 \exp \{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \}^2 y_{ij}} \right) + \gamma_{\phi 0}$$

e para a distribuição Lognormal por

$$\begin{aligned}
\pi(\phi|\mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}_{(*)}) &\propto \prod_{i=1}^n \prod_{j=1}^{n_i} \left[\exp \left\{ -\frac{(\log(y_{ij} \sqrt{\phi^{-1} + 1} / \exp \{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \}))^2}{2 \log(\phi^{-1} + 1)} \right\} \right]^{1 - \mathbf{I}_{\{y_{ij}=0\}}} \\
&\times \left[\frac{1}{y_{ij} \sqrt{2\pi \log(\phi^{-1} + 1)}} \right]^{1 - \mathbf{I}_{\{y_{ij}=0\}}} \phi^{\delta_{\phi 0} - 1} \exp \{ -\gamma_{\phi 0} \phi \},
\end{aligned} \tag{2.2.21}$$

cuja forma é conhecida a não ser por uma constante de proporcionalidade.

Para os efeitos aleatórios, temos que

$$\begin{aligned}
\pi(b_{ij}|\mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}) &\propto g(y_{ij}|\mu_{ij}, \phi, p_{ij}) \phi_{B_{ij}}(b_{ij}|\sigma_b^2) \\
&\propto f(y_{ij}|\mu_{ij}, \phi)^{1 - \mathbf{I}_{\{y_{ij}=0\}}} \phi_{B_{ij}}(b_{ij}|\sigma_b^2),
\end{aligned} \tag{2.2.22}$$

que depende da forma de $f(y_{ij}|\mu_{ij}, \phi)$ em cada distribuição, com $B_{ij} \stackrel{i.i.d}{\sim} N(0, \sigma_b^2)$.

Além disso, temos que

$$U_{2,i}|U_{1,i} = u_{1,i} \sim N(\psi u_{1,i}, \eta_2^2) \quad \text{e} \quad U_{1,i}|U_{2,i} = u_{2,i} \sim N\left(\frac{\psi \eta_1^2}{\psi^2 \eta_1^2 + \eta_2^2} u_{2,i}, \frac{\eta_1^2 - (\psi \eta_1^2)^2}{\psi^2 \eta_1^2 + \eta_2^2}\right)$$

e, assim,

$$\begin{aligned} \pi(u_{1,i}|\mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}) &\propto \prod_{i=1}^n \prod_{j=1}^{n_i} g(y_{ij}|\mu_{ij}, \phi, p_{ij}) \phi_{U_{1,i}|U_{2,i}=u_{2,i}}(u_{1,i}|\mathbf{\Sigma}_{\mathbf{u}}) \\ &\propto \prod_{i=1}^n \prod_{j=1}^{n_i} f(y_{ij}|\boldsymbol{\lambda}_{ij})^{1-\mathbf{I}_{\{y_{ij}=0\}}} \phi_{U_{1,i}|U_{2,i}=u_{2,i}}(u_{1,i}|\mathbf{\Sigma}_{\mathbf{u}}) \\ &= \prod_{i=1}^n \prod_{j=1}^{n_i} f(y_{ij}|\boldsymbol{\lambda}_{ij})^{1-\mathbf{I}_{\{y_{ij}=0\}}} \exp\left\{-\frac{\psi^2 \eta_1^2 + \eta_2^2}{2\eta_1^2 \eta_2^2} \left(u_{1,i} - \frac{\psi \eta_1^2}{\psi^2 \eta_1^2 + \eta_2^2} u_{2,i}\right)^2\right\} \\ \pi(u_{2,i}|\mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}) &\propto \prod_{i=1}^n \prod_{j=1}^{n_i} g(y_{ij}|\mu_{ij}, \phi, p_{ij}) \phi_{U_{2,i}|U_{1,i}=u_{1,i}}(u_{2,i}|\mathbf{\Sigma}_{\mathbf{u}}) \\ &\propto \prod_{i=1}^n \prod_{j=1}^{n_i} \left(\frac{p_{ij}}{1-p_{ij}}\right)^{\mathbf{I}_{\{y_{ij}=0\}}} (1-p_{ij}) \phi_{U_{2,i}|U_{1,i}=u_{1,i}}(u_{2,i}|\mathbf{\Sigma}_{\mathbf{u}}) \\ &= \prod_{i=1}^n \prod_{j=1}^{n_i} \left(\frac{p_{ij}}{1-p_{ij}}\right)^{\mathbf{I}_{\{y_{ij}=0\}}} (1-p_{ij}) \exp\left\{-\frac{1}{2\eta_2^2} (u_{2,i} - \psi u_{1,i})^2\right\}, \end{aligned} \tag{2.2.23}$$

onde $\phi_{U_{1,i}|U_{2,i}=u_{2,i}}(\cdot)$ e $\phi_{U_{2,i}|U_{1,i}=u_{1,i}}(\cdot)$, são as densidades de $U_{1,i}|U_{2,i} = u_{2,i}$ e $U_{2,i}|U_{1,i} = u_{1,i}$, respectivamente.

A condicional completa de $u_{1,i}$ também depende da forma de $f(y_{ij}|\mu_{ij}, \phi)$ em cada distribuição. Temos $\mu_{ij} = \exp\{\mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij}\}$, $p_{ij} = \exp\{\mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p + u_{2,i}\} / (1 + \exp\{\mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p + u_{2,i}\})$ e calculamos as condicionais para $u_{1,i}$ e b_{ij} em cada caso. Em todos os casos, não temos forma

fechada conhecida para as condicionais completas, ou seja, para a distribuição Gama

$$\begin{aligned} \pi(b_{ij}|\mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}) &\propto \left[(\exp\{b_{ij}\})^{-\phi} \exp \left\{ \frac{-\phi y_{ij}}{\exp\{\mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij}\}} \right\} \right]^{1-\mathbf{I}_{\{y_{ij}=0\}}} \\ &\times \exp \left\{ -\frac{1}{2\sigma_b^2} b_{ij}^2 \right\} \end{aligned} \quad (2.2.24)$$

$$\begin{aligned} \pi(u_{1,i}|\mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}) &\propto \prod_{i=1}^n \prod_{j=1}^{n_i} \left[(\exp\{u_{1,i}\})^{-\phi} \exp \left\{ \frac{-\phi y_{ij}}{\exp\{\mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij}\}} \right\} \right]^{1-\mathbf{I}_{\{y_{ij}=0\}}} \\ &\times \exp \left\{ -\frac{\psi^2 \eta_1^2 + \eta_2^2}{2\eta_1^2 \eta_2^2} \left(u_{1,i} - \frac{\psi \eta_1^2}{\psi^2 \eta_1^2 + \eta_2^2} u_{2,i} \right)^2 \right\}. \end{aligned}$$

Para a distribuição Inversa Gaussiana

$$\begin{aligned} \pi(b_{ij}|\mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}) &\propto \left[\exp \left\{ -\phi \frac{(y_{ij} - \exp\{\mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij}\})^2}{2 \exp\{\mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij}\}^2 y_{ij}} \right\} \right]^{1-\mathbf{I}_{\{y_{ij}=0\}}} \\ &\times \exp \left\{ -\frac{1}{2\sigma_b^2} b_{ij}^2 \right\} \end{aligned} \quad (2.2.25)$$

$$\begin{aligned} \pi(u_{1,i}|\mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}) &\propto \prod_{i=1}^n \prod_{j=1}^{n_i} \left[\exp \left\{ -\phi \frac{(y_{ij} - \exp\{\mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij}\})^2}{2 \exp\{\mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij}\}^2 y_{ij}} \right\} \right]^{1-\mathbf{I}_{\{y_{ij}=0\}}} \\ &\times \exp \left\{ -\frac{\psi^2 \eta_1^2 + \eta_2^2}{2\eta_1^2 \eta_2^2} \left(u_{1,i} - \frac{\psi \eta_1^2}{\psi^2 \eta_1^2 + \eta_2^2} u_{2,i} \right)^2 \right\}. \end{aligned}$$

Finalmente, para a distribuição Lognormal,

$$\begin{aligned}
\pi(b_{ij}|\mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}) &\propto \left[\exp \left\{ -\frac{\left(\log \left(y_{ij} \sqrt{\phi^{-1} + 1} / \exp \{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \} \right) \right)^2}{2 \log(\phi^{-1} + 1)} \right\} \right]^{1 - \mathbf{I}_{\{y_{ij}=0\}}} \\
&\quad \times \exp \left\{ -\frac{1}{2\sigma_b^2} b_{ij}^2 \right\} \\
\pi(u_{1,i}|\mathbf{H}, \mathbf{y}, \mathbf{u}, \mathbf{X}_\mu, \mathbf{X}_p, \boldsymbol{\theta}) &\propto \prod_{i=1}^n \prod_{j=1}^{n_i} \left[\exp \left\{ -\frac{\left(\log \left(y_{ij} \sqrt{\phi^{-1} + 1} / \exp \{ \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \} \right) \right)^2}{2 \log(\phi^{-1} + 1)} \right\} \right]^{1 - \mathbf{I}_{\{y_{ij}=0\}}} \\
&\quad \times \exp \left\{ -\frac{\psi^2 \eta_1^2 + \eta_2^2}{2\eta_1^2 \eta_2^2} \left(u_{1,i} - \frac{\psi \eta_1^2}{\psi^2 \eta_1^2 + \eta_2^2} u_{2,i} \right)^2 \right\}.
\end{aligned} \tag{2.2.26}$$

2.3 Critérios de comparação de modelos

Neste trabalho, consideramos dois critérios de comparação de modelos: um critério baseado no logaritmo da pseudo verossimilhança marginal (LPML) derivado a partir da ordenada preditiva condicional (CPO), proposta por Geisser e Eddy (1979) e por Geisser (1993), e um critério a partir da função desvio (DIC) adaptado para o caso de presença de efeitos aleatórios (Celeux et al., 2006).

Seja $\mathbf{y} = (y_1, \dots, y_n)$ a amostra independente observada e $\mathbf{y}_{(i)}$ a amostra sem a observação i , $i = 1, \dots, n$. O CPO associado à observação i é definido como a densidade marginal preditiva a posteriori de y_i condicional às observações restantes em $\mathbf{y}_{(i)}$ e é calculado como

$$\text{CPO}_i = f(y_i|\mathbf{y}_{(i)}) = \int f(y_i|\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\mathbf{y}_{(i)})d\boldsymbol{\theta}, \tag{2.3.1}$$

onde $\boldsymbol{\theta}$ é o vetor de parâmetros desconhecidos, $f(y_i|\boldsymbol{\theta})$ é a distribuição condicional de y_i dado $\boldsymbol{\theta}$ e $\pi(\boldsymbol{\theta}|\mathbf{y}_{(i)})$ é a distribuição a posteriori associada ao vetor $\mathbf{y}_{(i)}$.

A intuição por trás do CPO é comparar os modelos através das densidades preditivas $f(y_i|\mathbf{y}_{(i)})$ via validação cruzada (Gelfand, 1996). As distribuições preditivas das observações que foram removidas são obtidas para examinar se os pontos removidos pertencem à regiões com densidade considerável. Assim, o modelo com maior poder de predição em termos da densidade preditiva a posteriori é aquele onde, em média, os pontos removidos tem maior chance

de serem “explicados”. O CPO pode ser visto como uma medida do efeito da observação y_i na densidade preditiva a priori de \mathbf{y} , $f(\mathbf{y}) = \int f(\mathbf{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) d\boldsymbol{\theta}$, uma vez que

$$\text{CPO}_i = f(y_i|\mathbf{y}_{(i)}) = \frac{f(\mathbf{y})}{f(\mathbf{y}_{(i)})}. \quad (2.3.2)$$

Em geral, a forma de (2.3.1) não pode ser reconhecida. Porém, note que

$$\begin{aligned} (\text{CPO}_i)^{-1} &= \frac{1}{f(y_i|\mathbf{y}_{(i)})} = \frac{f(\mathbf{y}_{(i)})}{f(\mathbf{y})} = \int \frac{f(\mathbf{y}_{(i)}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{f(\mathbf{y})} d\boldsymbol{\theta} = \int \frac{f(y_i|\boldsymbol{\theta})}{f(y_i|\mathbf{y})} \frac{f(\mathbf{y}_{(i)}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{f(\mathbf{y})} d\boldsymbol{\theta} \\ &= \int \frac{1}{f(y_i|\boldsymbol{\theta})} \frac{f(\mathbf{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{f(\mathbf{y})} d\boldsymbol{\theta} = \int \frac{1}{f(y_i|\boldsymbol{\theta})} \pi(\boldsymbol{\theta}|\mathbf{y}) d\boldsymbol{\theta} = E_{\boldsymbol{\theta}|\mathbf{y}} \left[\frac{1}{f(y_i|\boldsymbol{\theta})} \right]. \end{aligned} \quad (2.3.3)$$

Dey et al. (1997) mostram que uma estimativa por Monte Carlo de (2.3.1) através da aproximação por média harmônica de (2.3.3) é dada por

$$\widehat{\text{CPO}}_i = \left(\frac{1}{M} \sum_{j=1}^M \frac{1}{f(y_i|\boldsymbol{\theta}^{(j)})} \right)^{-1}, \quad (2.3.4)$$

onde $\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(M)}$ são amostras da posteriori $\pi(\boldsymbol{\theta}|\mathbf{y}_{(i)})$. Assim, temos uma estimativa que não necessita que o modelo seja reajustado a cada observação retirada.

O produto dos CPOs individuais pode ser visto como uma pseudo verossimilhança marginal (Carlin e Louis, 2008), com valores altos indicando um bom ajuste geral do modelo. Então, consideramos a estatística LPML como resumo dos CPOs tomando o logaritmo do produto, ou seja, $\text{LPML} = \sum_{i=1}^n \log(\text{CPO}_i)$, com valores mais altos indicando melhor ajuste.

Raftery et al. (2006) mostram que estimadores aproximados através da média harmônica para a verossimilhança marginal $f(\mathbf{y})$ podem ter variância infinita e sugerem integrar a verossimilhança em um subconjunto do espaço paramétrico como uma tentativa de estabilizar o estimador. Lewis et al. (2014) dão uma intuição de que, embora os estimadores da verossimilhança marginal calculados através da média harmônica sejam instáveis, os estimadores para os CPOs individuais devem ser menos instáveis. Congdon (2005) sugere que se o logaritmo da verossimilhança individual ultrapassar -10 ou -20, podemos considerar o estimador por média harmônica.

No caso de modelos com efeitos aleatórios o cálculo das verossimilhanças $f(\mathbf{y}_i|\boldsymbol{\theta}^{(j)})$ envolve calcular uma integral com respeito aos efeitos aleatórios. Essa integral podem trazer uma maior estabilidade do estimador (2.3.4) uma vez que o espaço paramétrico é reduzido (Raftery et al., 2006). Neste trabalho, o espaço paramétrico $\boldsymbol{\theta}$ é aumentado por \mathbf{u} , como na Seção 2.2, como sugerido em Sinharay e Stern (2005) para o cálculo do fator de Bayes quando os efeitos aleatórios são de grande dimensão, tornando as integrais computacionalmente dispendiosas. Assim, para modelos com efeitos aleatórios para respostas positivas aumentadas em zeros em (1.1.40), temos que a estimativa dos CPOs é dada por

$$\widehat{\text{CPO}}_i = \left(\frac{1}{M} \sum_{j=1}^M \frac{1}{g(y_i|\mathbf{u}_i^{(j)}, \boldsymbol{\theta}^{(j)})} \right)^{-1}, \quad (2.3.5)$$

onde $\mathbf{u}_i^{(j)}$ e $\boldsymbol{\theta}^{(j)}$ são amostras obtidas a partir da posteriori.

O critério DIC foi proposto por Spiegelhalter et al. (2002) como critério de comparação de modelos considerando o ajuste e a complexidade do modelo no seu cálculo. No caso dos modelos somente com efeitos fixos, essa complexidade é dada pelo número efetivo de parâmetros p_D . No caso de modelos com efeitos aleatórios essa complexidade demanda uma alternativa ao DIC tradicional (Celeux et al., 2006).

O DIC é definido por $\text{DIC} = \overline{D(\boldsymbol{\theta})} + p_D$, onde $\overline{D(\boldsymbol{\theta})} = E[D(\boldsymbol{\theta})|\mathbf{y}]$ é a esperança a posteriori da função desvio do modelo $D(\boldsymbol{\theta})$, dada por $-2\log(f(\mathbf{y}|\boldsymbol{\theta}))$, que é uma medida do ajuste do modelo; $p_D = \overline{D(\boldsymbol{\theta})} - D(\tilde{\boldsymbol{\theta}}) = E[D(\boldsymbol{\theta})|\mathbf{y}] - D(\tilde{\boldsymbol{\theta}})$ é a diferença da esperança a posteriori da função desvio e da função desvio avaliada em $\tilde{\boldsymbol{\theta}}$ que é uma estimativa de $\boldsymbol{\theta}$ que depende de \mathbf{y} , e pode ser tomada como uma penalidade devido à complexidade do modelo. Assim, o DIC pode ser escrito como

$$\text{DIC} = -4E[\log(f(\mathbf{y}|\boldsymbol{\theta}))|\mathbf{y}] + 2\log(f(\mathbf{y}|\tilde{\boldsymbol{\theta}})), \quad (2.3.6)$$

onde $\tilde{\boldsymbol{\theta}}$ é uma estimativa de $\boldsymbol{\theta}$.

No caso do modelo com efeitos aleatórios para respostas positivas aumentadas em zeros, consideraremos o critério DIC_7 proposto em Celeux et al. (2006), o qual calcula $\overline{D(\boldsymbol{\theta})}$ e $D(\tilde{\boldsymbol{\theta}})$ condicionalmente aos efeitos aleatórios, como utilizado em Neelon et al. (2010) para modelos com zero inflacionados e efeitos aleatórios, onde a integral com respeito aos efeitos aleatórios não é calculada. Celeux et al. (2006) calcula o DIC a partir da verossimilhança marginal e

o chama de DIC_3 , apontando que o DIC_3 é mais confiável na comparação de modelos do que critérios condicionais como o DIC_7 . Porém, neste trabalho levamos em consideração as dificuldades de calcular as integrais nos efeitos aleatórios quando esses tem grande dimensão, como no caso do modelo proposto em (1.1.40).

O critério DIC_7 para o modelo em (1.1.40) é baseado na verossimilhança condicional $g(\mathbf{y}|\mathbf{u}, \boldsymbol{\theta})$ e é dado por

$$DIC_7 = -4E[\log(g(\mathbf{y}|\mathbf{u}, \boldsymbol{\theta}))|\mathbf{y}] + 2\log(g(\mathbf{y}|\hat{\mathbf{u}}, \hat{\boldsymbol{\theta}})), \quad (2.3.7)$$

onde $\hat{\mathbf{u}}$ é a média a posteriori dos efeitos aleatórios e $\hat{\boldsymbol{\theta}}$ é a média a posteriori de $\boldsymbol{\theta}$ obtidas através das amostras das respectivas distribuições a posteriori.

Aproximamos a esperança a posteriori da função desvio $\overline{D(\boldsymbol{\theta}, \mathbf{u})} = E[\log(g(\mathbf{y}|\mathbf{u}, \boldsymbol{\theta}))|\mathbf{y}]$ por $(1/M) \sum_{j=1}^M \sum_{i=1}^n \log(g(\mathbf{y}_i|\mathbf{u}_i^{(j)}, \boldsymbol{\theta}^{(j)}))$ e calculamos $D(\hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}) = \log(g(\mathbf{y}|\hat{\mathbf{u}}, \hat{\boldsymbol{\theta}}))$ avaliando o logaritmo da função $g(\mathbf{y}|\cdot)$ nas médias $\hat{\mathbf{u}}$ e $\hat{\boldsymbol{\theta}}$, dos efeitos aleatórios $\mathbf{u}^{(j)}$ e dos parâmetros $\boldsymbol{\theta}^{(j)}$, respectivamente, onde $\mathbf{u}^{(j)}$ e $\boldsymbol{\theta}^{(j)}$ são amostras obtidas da posteriori. Assim como para o DIC, considera-se o melhor modelo aquele com menor DIC_7 .

2.4 Análise de observações influentes

Uma observação pode ser considerada influente se sua exclusão produz diferenças significativas no ajuste do modelo, como por exemplo nas estimativas dos parâmetros e o respectivo erro padrão. Uma maneira de determinarmos se uma observação é influente no ajuste do modelo é acessar a qualidade do ajuste ao deletarmos esta observação como nos métodos propostos por Cook e Weisberg (1982).

Seja $\pi(\boldsymbol{\theta}|\mathbf{y})$ a distribuição a posteriori de $\boldsymbol{\theta}$ considerando $\mathbf{y} = (y_1, \dots, y_n)$ e seja $\pi(\boldsymbol{\theta}|\mathbf{y}_{(i)})$ a distribuição a posteriori de $\boldsymbol{\theta}$ onde $\mathbf{y}_{(i)}$ é o vetor \mathbf{y} sem a observação y_i . A medida da divergência $d_q(i)$ entre $\pi(\boldsymbol{\theta}|\mathbf{y})$ e $\pi(\boldsymbol{\theta}|\mathbf{y}_{(i)})$ para a observação i é dada por

$$d_q(i) = \int q \left(\frac{\pi(\boldsymbol{\theta}|\mathbf{y}_{(i)})}{\pi(\boldsymbol{\theta}|\mathbf{y})} \right) \pi(\boldsymbol{\theta}|\mathbf{y}) d\boldsymbol{\theta}, \quad (2.4.1)$$

onde $q(\cdot)$ é uma função convexa tal que $q(1) = 0$. A divergência $d_q(i)$ pode ser interpretada como a divergência q do efeito na posteriori de $\boldsymbol{\theta}$ ao deletarmos a observação y_i .

Diferentes medidas de divergência são definidas a partir da forma de $q(\cdot)$. A divergência de Kullback-Leibler considera $q(z) = -\log(z)$; a J-distância é obtida para $q(z) = (z - 1)\log(z)$ e a L_1 -distância toma $q(z) = |z - 1|$.

Utilizando a estatística CPO_i em (2.3.2) e multiplicando o numerador e o denominador por $f(y_i|\boldsymbol{\theta})$, temos que

$$\frac{\pi(\boldsymbol{\theta}|\mathbf{y}_{(i)})}{\pi(\boldsymbol{\theta}|\mathbf{y})} = \frac{\pi(\boldsymbol{\theta})f(\mathbf{y}_{(i)}|\boldsymbol{\theta})}{f(\mathbf{y}_{(i)})} \frac{f(\mathbf{y})}{\pi(\boldsymbol{\theta})f(\mathbf{y}|\boldsymbol{\theta})} = \text{CPO}_i \frac{f(\mathbf{y}_{(i)}|\boldsymbol{\theta})}{f(\mathbf{y}|\boldsymbol{\theta})} = \text{CPO}_i \frac{f(\mathbf{y}_{(i)}|\boldsymbol{\theta})}{f(\mathbf{y}|\boldsymbol{\theta})} \times \frac{f(y_i|\boldsymbol{\theta})}{f(y_i|\boldsymbol{\theta})} = \frac{\text{CPO}_i}{f(y_i|\boldsymbol{\theta})}. \quad (2.4.2)$$

Logo, podemos reescrever a divergência q em (2.4.1) como $d_q(i) = E_{\boldsymbol{\theta}|\mathbf{y}} [q(\text{CPO}_i/f(\mathbf{y}_i|\boldsymbol{\theta}))]$, onde a esperança é calculada a partir da posteriori $\pi(\boldsymbol{\theta}|\mathbf{y})$. Assim, uma estimativa por Monte-Carlo de $d_q(i)$ é dada por:

$$\widehat{d_q(i)} = \frac{1}{M} \sum_{j=1}^M q \left(\frac{\widehat{\text{CPO}}_i}{f(y_i|\boldsymbol{\theta}^{(j)})} \right), \quad (2.4.3)$$

onde $\boldsymbol{\theta}^{(j)}$ são amostras obtidas da posteriori.

Para determinar o ponto de corte que considera as observações como influentes, consideramos os pontos de corte como em Rodrigues-Motta et al. (2015), ou seja, uma observação é influente se $d_{L_1}(i) \geq 0.80$ para a L_1 -distância, se $d_{K-L}(i) \geq 0.51$ para a divergência de Kullback-Leibler e se $d_J(i) \geq 0.88$ para a J-distância.

Em nosso estudo, consideramos a estimativa dos CPOs como em (2.3.5) e, assim, temos que (2.4.3) é dada por

$$\widehat{d_q(i)} = \frac{1}{M} \sum_{j=1}^M q \left(\frac{\widehat{\text{CPO}}_i}{g(y_i|\mathbf{u}_i^{(j)}, \boldsymbol{\theta}^{(j)})} \right). \quad (2.4.4)$$

Capítulo 3

Estudo de simulação sobre a distribuição a priori das componentes de dispersão dos efeitos aleatórios

3.1 Introdução

Lambert et al. (2005) observam que, quando temos uma quantidade pequena de medidas contribuindo para estimar variâncias e covariâncias entre unidades experimentais, inferências sobre a estimativa da matriz de covariância podem ser sensíveis à escolha da priori para a matriz de covariância dos efeitos aleatórios. Browne e Draper (2000) observaram esse problema utilizando a distribuição Wishart como priori. Diante disso, propomos a partir da literatura quatro tipos de priori para a matriz de covariância Σ definida em (1.1.43), dadas a seguir.

(i) Priori Tipo I

Assumimos que

$$\Sigma_{\mathbf{u}}^{-1} \sim \text{Wishart}(V, \nu_u) \quad \text{e} \quad \sigma_b^2 \sim \text{Inv-Gama}(\delta_{b0}, \gamma_{b0}), \quad (3.1.1)$$

com V , ν_u , δ_{b0} , γ_{b0} conhecidos.

(iii) Priori Tipo II

Seguindo Gelman et al. (2006), assumimos

$$\sigma_{u_1}, \sigma_{u_2}, \sigma_b \sim U(0, A_0) \quad \text{e} \quad \rho \sim U(-1, 1), \quad (3.1.2)$$

onde $U(\cdot, \cdot)$ representa a distribuição Uniforme e A_0 é escolhido de forma que o intervalo seja grande.

(iv) Priori Tipo III

Seguindo Chung et al. (2015), propomos

$$\sigma_{u_1}^2, \sigma_{u_2}^2, \sigma_b^2 \sim \text{Gama}(1.5, \theta_0) \quad \text{e} \quad \frac{\rho + 1}{2} \sim \text{Beta}(1.5, 1.5), \quad (3.1.3)$$

para θ_0 pequeno e conhecido.

(ii) Priori Tipo IV

Seguindo Neelon et al. (2010), temos que

$$U_{1,i} \sim N(0, \eta_1^2) \quad \text{e} \quad U_{2,i}|U_{1,i} = u_{1,i} \sim N(\psi u_{1,i}, \eta_2^2), \quad (3.1.4)$$

tal que a matriz $\Sigma_{\mathbf{u}}$ passa a ser escrita como

$$\Sigma_{\mathbf{u}} = \begin{pmatrix} \eta_1^2 & \psi \eta_1^2 \\ \psi \eta_1^2 & \psi^2 \eta_1^2 + \eta_2^2 \end{pmatrix}, \quad (3.1.5)$$

onde

$$\sigma_{u_1}^2 = \eta_1^2, \quad \sigma_{u_2}^2 = \psi^2 \eta_1^2 + \eta_2^2, \quad \rho = \frac{\psi \eta_1^2}{\sqrt{\eta_1^2(\psi^2 \eta_1^2 + \eta_2^2)}}, \quad (3.1.6)$$

com $\psi \sim N(0, \sigma_{\psi 0}^2)$ e $\eta_1^2, \eta_2^2 \sim \text{Inv-Gama}(\delta_{\eta 0}, \gamma_{\eta 0})$, com hiperparâmetros conhecidos. Além disso, $\sigma_b^2 \sim \text{Inv-Gama}(\delta_{b0}, \gamma_{b0})$.

3.2 Estudo de simulação da distribuição a priori das componentes de variância a partir do modelo de regressão Gama aumentada em zero

Nesta seção, fazemos um estudo de simulação a partir das prioris tipo I, II, III, IV descritas na Seção 3.1, considerando (1.1.45) com $f(y_{ij}|\mu_{ij}, \phi)$ sendo representada por uma distribuição Gama com parâmetros μ_{ij} e ϕ .

O estudo utiliza amostras de tamanho $n = 50, 150$ e considera 100 simulações para cada n . Tomamos em cada caso, 2 observações (visitas) para cada unidade amostral ($j = 1, 2$).

Assim, geramos $\mathbf{y}_1, \dots, \mathbf{y}_n$ vetores aleatórios e independentes, tal que, condicionado em $\mathbf{U}_i = \mathbf{u}_i$, $y_{ij}|\mathbf{u}_i$ é a realização de $Y_{ij}|\mathbf{u}_i \sim g(y_{ij}|\mu_{ij}, \phi, p_{ij})$ e \mathbf{u}_i é a realização da variável aleatória \mathbf{U}_i tal que $\mathbf{U}_i \sim N(0, \Sigma)$ com Σ definida em (1.1.43). Neste estudo, consideramos $\sigma_{u_1} = 0, 1$, $\sigma_{u_2} = 0, 3$ e $\sigma_b = 0, 3$. Além disso, consideramos ρ igual a $-0, 9, -0, 5, 0, 0, 1, 0, 5$ e $0, 9$. Também, construímos

$$\begin{aligned} \log(\mu_{ij}) &= \alpha_0 + \alpha_1 x_{ij} + u_{1,i} + b_{ij} \\ \text{logito}(p_{ij}) &= \beta_0 + \beta_1 z_{ij} + u_{2,i} \end{aligned} \quad , \quad (3.2.1)$$

considerando $\alpha_0 = 2$, $\alpha_1 = 1$, $\beta_0 = -2$, $\beta_1 = -1$. Aqui, $x_{ij} = x_i$ foi gerado de uma distribuição Bernoulli de parâmetro 0,8 e $z_{ij} = z_i$ foi gerado de uma distribuição Bernoulli de parâmetro 0,8, independente de x_{ij} . Finalmente, tomamos $\phi = 3$.

Para as prioris dos efeitos fixos consideramos $\alpha_0, \alpha_1, \beta_0, \beta_1 \sim N(0, 10^3)$. Para a priori do parâmetro de dispersão tomamos $\phi \sim \text{Gama}(0,01, 0,01)$. Para os parâmetros de dispersão, em cada priori, temos:

(i) Priori Tipo I

$$\Sigma_u^{-1} \sim \text{Wishart}(0, 1 \mathbf{I}_2, 2) \quad \text{e} \quad \sigma_b^2 \sim \text{Inv-Gama}(0,01, 0,01). \quad (3.2.2)$$

(iii) Priori Tipo II

$$\sigma_{u_1}, \sigma_{u_2}, \sigma_b \sim U(0, 100) \quad \text{e} \quad \rho \sim U(-1, 1). \quad (3.2.3)$$

(iv) Priori Tipo III

$$\sigma_{u_1}^2, \sigma_{u_2}^2, \sigma_b^2 \sim \text{Gama}(1, 5, 0, 01) \quad \text{e} \quad \frac{\rho + 1}{2} \sim \text{Beta}(1, 5, 1, 5). \quad (3.2.4)$$

(ii) Priori Tipo IV

$$u_{1,i} \sim N(0, \eta_1^2) \quad \text{e} \quad u_{2,i}|u_{1,i} \sim N(\psi u_{1,i}, \eta_2^2), \quad (3.2.5)$$

com $\psi \sim N(0, 10^3)$ e $\eta_1^2, \eta_2^2 \sim \text{Inv-Gama}(0, 01, 0, 01)$. Além disso, $\sigma_b^2 \sim \text{Inv-Gama}(0, 01, 0, 01)$.

Quando $\rho = 0$, temos que as prioris Tipo I e IV coincidem, tal que $\sigma_{u_1}^2$ e $\sigma_{u_2}^2$ tem distribuição $\text{Inv-Gama}(0, 01, 0, 01)$.

O objetivo é amostrar de (2.2.2) considerando os diferentes cenários e verificar a performance do ajuste a partir das prioris Tipo I, II, III e IV descritas na Seção 3.1.

Para cada ajuste do modelo, foram usadas duas cadeias Markov independentes de tamanho 100.000 cada começando em valores distantes, com aquecimento inicial de 50.000 pontos e espaçamento de tamanho 100, resultando em 1.000 observações da posteriori.

Utilizamos três medidas para comparar o processo de estimação do modelo utilizando cada distribuição a priori ao longo dos cenários: o erro quadrático médio (EQM), o erro relativo (ER) e a proporção de vezes que o intervalo de credibilidade contém o valor verdadeiro do parâmetro, ou a probabilidade de cobertura (PC), apresentadas a seguir. Seja $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_0, \beta_1, \phi, \sigma_{u_1}^2, \sigma_{u_2}^2, \rho, \sigma_b^2)$ o vetor de parâmetros com os verdadeiros valores e seja θ^* um elemento de $\boldsymbol{\theta}$. Seja $\hat{\theta}_k$ a média a posteriori obtida na amostra k para θ^* , $k = 1, \dots, 100$. Assim,

$$\begin{aligned} \text{EQM}(\hat{\theta}^*) &= \frac{1}{100} \sum_{k=1}^{100} (\hat{\theta}_k - \theta^*)^2 \\ \text{ER}(\hat{\theta}^*) &= \frac{1}{100} \sum_{k=1}^{100} \left(\frac{\hat{\theta}_k}{\theta^*} - 1 \right) \\ \text{PC}(\hat{\theta}^*) &= \frac{1}{100} \sum_{k=1}^{100} I(\theta^* \in [\hat{\theta}_{k,I}, \hat{\theta}_{k,S}]) \end{aligned} \quad (3.2.6)$$

onde $I(\theta^* \in [\hat{\theta}_{k,I}, \hat{\theta}_{k,S}])$ é uma função indicadora que é igual a um se θ^* está no intervalo

$[\hat{\theta}_{k,I}, \hat{\theta}_{k,S}]$ e zero caso contrário. Aqui, $\hat{\theta}_{k,I}$ e $\hat{\theta}_{k,S}$ representam os percentis amostrais 2,5% e 97,5% da posteriori, respectivamente. Para essas medidas, temos que os melhores estimadores são aqueles que tem erro relativo e erro quadrático médio próximos de zero e probabilidade de cobertura próxima de um.

Os resultados para cada cenário de simulação são apresentados no Apêndice A a partir das Tabelas A.1, A.2, A.3, A.4, A.5 e A.6.

Para $\rho = 0$,

$$\Sigma_{\mathbf{u}} = \begin{pmatrix} \sigma_{u_1}^2 & 0 \\ 0 & \sigma_{u_2}^2 \end{pmatrix},$$

e assim, as prioris do Tipo I e IV produzem o mesmo resultado. A partir da Tabela A.1, notamos que as prioris Tipo I e Tipo IV produzem os melhores resultados, já que o ER, EQM são em geral mais próximos de zero e os valores para PC próximos de um.

Nas Tabelas A.2 a A.6, encontramos os valores de ER, EQM e PC para $\rho > 0$ e $\rho < 0$ quando $n = 50$ e $n = 150$. Em geral, a priori do Tipo IV apresenta o melhor desempenho com respeito ao ER, EQM e à PC, sendo que para o valor de $n = 150$ os valores são ainda mais satisfatórios. Para alguns casos, a priori Tipo II apresenta probabilidade de cobertura maior do que a priori Tipo IV. No entanto, o ER e o EQM em geral são melhores para a priori Tipo IV. Entre as propostas, a priori Tipo III apresenta o pior desempenho para estimar o parâmetro ϕ , independentemente do valor de n . A probabilidade de cobertura de β e ϕ para as prioris do Tipo I e III é em geral mais baixa do que as outras prioris quando $\rho > 0$ e $\rho < 0$. Em geral, as medidas indicam melhor desempenho quando o número de observações é maior. O problema de probabilidade de baixa cobertura foi abordado por Lambert et al. (2005) para o caso em que as componentes de dispersão têm valor próximo de zero, onde as estimativas obtidas tendem a ser viesadas para cima, uma vez que apenas valores positivos são permitidos. Assim, nesse estudo, optamos pela priori do Tipo IV.

3.3 Estudo de simulação a partir do número de observações por unidade amostral para o modelo Gama aumentada em zero

Considerando a priori do Tipo IV, estamos interessados em saber se obtemos melhores estimativas para os parâmetros do modelo quando aumentamos o número de observações por unidade amostral.

Neste estudo de simulação, consideramos o esquema proposto na Seção 3.2 e tomamos o número de visitas igual a 5, 10 e 15. Comparamos os resultados utilizando novamente o erro relativo, o erro quadrático médio e a probabilidade de cobertura definidos em (3.2.6) a partir de 50 amostras de tamanhos amostrais $n = 50, 150$.

Realizamos a simulação para valores ρ iguais a $-0,9, -0,5, 0,1, 0,5$ e $0,9$. Os ajustes foram realizados como descrito no estudo de simulação na Seção 3.2. Os resultados são apresentados nas Tabelas B.1, B.2, B.3, B.4 e B.5 do Apêndice B.

Observamos uma melhora nas estimativas dos parâmetros quando aumentamos o número de visitas, mas a melhora é mais significativa quando aumentamos o tamanho amostral. Por exemplo, para todos os cenários, se compararmos o ER, o EQM e a PC quando temos 5 visitas e tamanho amostral igual a 150, temos melhores resultados do que quando temos 10 visitas e tamanho amostral igual a 50. Além disso, parece haver uma combinação entre o tamanho amostral e o número de visitas, já que, por exemplo, para σ_{u_1} , obtemos valores menores do ER e EQM e maiores da PC quando o tamanho amostral é 50 com 15 visitas comparando com o tamanho amostral 150 com 5 visitas. No geral, as medidas indicam que o ganho ao se aumentar o número de visitas é substancial nas estimativas dos parâmetros σ_{u_2} , ρ e ϕ .

Capítulo 4

Aplicação

Nesse Capítulo apresentamos a aplicação dos métodos e resultados a partir de dados reais sobre consumo habitual. Os dados estudados nesse Capítulo foram cedidos pelo Prof. Dr. Eli-seu Verly Junior da Universidade Estadual do Rio de Janeiro (UERJ) e coordenador do projeto “Variabilidade intrapessoal do consumo dietético e viés em análises de estudos epidemiológicos” com apoio do CNPq (481434/2013-5) e da FAPERJ (E-26/111.822/2012).

Estudos sobre consumo alimentar têm sido utilizados para avaliar hábitos alimentares de indivíduos de uma população e investigar a sua adequação à padrões de dieta recomendados, e avaliar a relação do consumo de certos alimentos com estado de saúde para certas doenças (Kipnis et al., 2009). Para acompanhar os padrões alimentares, deseja-se estudar o consumo habitual, que é definido como o consumo a longo prazo de certo alimento ou nutriente.

Nesse estudo, para estimar a distribuição do consumo habitual de indivíduos do Rio de Janeiro foram utilizados os inquéritos recordatórios de 24 horas (Dodd et al., 2006), onde os participantes reportam o que comeram e beberam nas 24 horas anteriores à entrevista, tomada em dias não consecutivos. As informações são então repassadas para um software construído para pesquisas alimentares brasileiras que contém padrões de porções e receitas.

Um dos objetivos do estudo é modelar a distribuição do consumo habitual de frutas legumes e vegetais (em gramas) dos indivíduos dessa população. Uma vez que o consumo assume valores num intervalo contínuo positivo ou o valor zero, utilizaremos a metodologia dos modelos mistos para respostas positivas aumentadas de zeros apresentados em Rodrigues-Motta et al. (2015) estudados neste trabalho.

Por utilizar essa classe de modelos, evitamos o uso de transformações visando a normalidade das respostas positivas, pois essas em geral apresentam distribuição assimétrica. Devido à essa

assimetria, utilizamos as distribuições Normal, Inversa Gaussiana e Lognormal para acomodar as respostas positivas. Além disso, modelamos a probabilidade de obter valor zero e a média das respostas positivas através das covariáveis presentes no estudo e de efeitos aleatórios, que buscam acomodar tanto a variabilidade entre as respostas de um mesmo indivíduo como a variabilidade entre os diferentes indivíduos. A aplicação é realizada utilizando a metodologia sob a abordagem Bayesiana estudada. Os modelos foram ajustados utilizando o software **OpenBUGS** (Thomas et al., 2006) pelo software **R** através da biblioteca **BRugs** (Ligges et al., 2009).

4.1 Análise descritiva

Neste trabalho, estudamos os dados provenientes de um estudo sobre consumo alimentar conduzido na cidade do Rio de Janeiro entre Março de 2013 e Abril de 2014. Nesse estudo, foi obtido o consumo (em gramas) de frutas, legumes e vegetais (FLV) de 162 indivíduos através de entrevistas realizadas em 19 dias não consecutivos. As entrevistas, chamadas de visitas nesse contexto, foram realizadas em diferentes dias da semana e não foram administradas em um único dia da semana repetidamente para um mesmo indivíduo. Além disso, foram observadas características sociais e demográficas como idade, IMC (Índice de Massa Corporal), sexo, consumo de refrigerante, se possui consumo adequado de FLV (opinião), renda e escolaridade. A descrição das variáveis do estudo encontra-se na Tabela 4.1. Nesse estudo, o valor das características sociais e demográficas é constante ao longo do estudo para cada indivíduo. O período médio de acompanhamento de cada indivíduo foi igual a 3 meses.

Tabela 4.1: Descrição das variáveis presentes no estudo.

Variável	Descrição	Categorias
Idade	Idade em anos	
IMC	Índice de Massa Corporal	
Sexo	Sexo do entrevistado	1- Masculino 2- Feminino
Refrigerante	Consumo alto de refrigerante	1- Sim 2- Não
FLV Adequado	Consumo adequado (opinião)	1- Sim 2- Não
Renda	Renda familiar em reais	1- < 2.000 2- 2.000-4.000 3- 4.000-6.000 4- > 6.000
Escolaridade	Escolaridade máxima	1- Fundamental 2- Médio 3- Superior 4- Outro
Peso	Peso antes do estudo	0- Igual 1- Aumentou 2- Diminuiu
Consumo	Consumo de FLV em gramas	

A assimetria à direita na distribuição amostral do consumo positivo pode ser notada pelos valores de assimetria amostral maiores que zero na Tabela 4.2 e pelos gráficos Box-plot na Figura 4.1 ao longo das visitas. As variação das respostas positivas entre os indivíduos parece ser significativa dentro de uma mesma visita. Além disso, podemos notar que temos porcentagens

de zeros acima de 15% para praticamente todas as visitas. Observando as frequências da quantidade de dias em que cada indivíduo teve consumo zero durante o estudo apresentadas na Figura 4.2, podemos observar que a maioria dos indivíduos possui poucos dias de consumo igual a zero individualmente, de modo que muitos indivíduos apresentam nenhum dia ou apenas um ou dois dias sem consumo positivo.

Tabela 4.2: Porcentagem de zeros e medidas-resumo para os valores positivos do consumo de FLV em cada visita.

Visita	Zeros (%)	Média	DP	CV	Assimetria	Q ₁	Mediana	Q ₃
1	16,05	197,56	150,40	0,76	0,83	75,00	150,00	290,00
2	11,73	198,70	155,88	0,78	1,13	71,22	152,00	287,50
3	12,35	206,83	180,48	0,87	1,52	74,58	144,50	285,44
4	15,43	204,36	206,99	1,01	2,30	69,88	142,88	284,44
5	16,67	204,61	200,23	0,98	2,16	67,22	146,44	286,50
6	19,14	217,70	183,52	0,84	1,04	71,26	175,00	360,61
7	16,67	186,09	176,68	0,95	1,93	61,49	144,44	236,50
8	22,22	191,28	153,32	0,80	1,22	81,00	146,66	282,47
9	16,67	212,07	199,81	0,94	1,65	65,00	160,00	290,00
10	17,28	218,09	180,37	0,83	1,72	80,00	181,94	286,50
11	18,52	188,46	163,00	0,86	1,68	75,00	132,50	259,16
12	16,67	211,02	179,60	0,85	1,45	69,44	162,00	292,08
13	18,52	190,99	176,03	0,92	2,01	63,75	150,00	255,61
14	15,43	184,72	195,27	1,06	2,93	65,00	123,32	240,00
15	22,84	215,17	154,14	0,72	0,94	93,50	183,32	300,00
16	16,05	191,26	189,27	0,99	1,96	60,75	119,94	270,00
17	19,14	211,79	241,05	1,14	3,25	64,72	148,88	263,72
18	15,43	197,41	198,76	1,01	1,63	54,00	130,00	239,00
19	19,14	187,13	181,26	0,97	2,16	68,50	130,00	239,94

4.2 Análise Inferencial

Para o ajuste dos modelos mistos para respostas positivas aumentadas em zero apresentados em (1.1.40) aos dados utilizamos as distribuições Gama, Inversa Gaussiana e Lognormal para acomodar os valores positivos. Aqui, y_{ij} é o consumo observado de frutas, legumes e vegetais (em gramas) na j -ésima visita da entrevista associada ao indivíduo i , onde $i = 1, \dots, 162$ e $j = 1, \dots, 19$. Modelamos neste estudo a média μ_{ij} e a probabilidade de zero p_{ij} , através de

$$\begin{aligned} \log(\mu_{ij}) &= \mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu + u_{1,i} + b_{ij} \\ \text{logito}(p_{ij}) &= \mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p + u_{2,i} \end{aligned}, \quad (4.2.1)$$

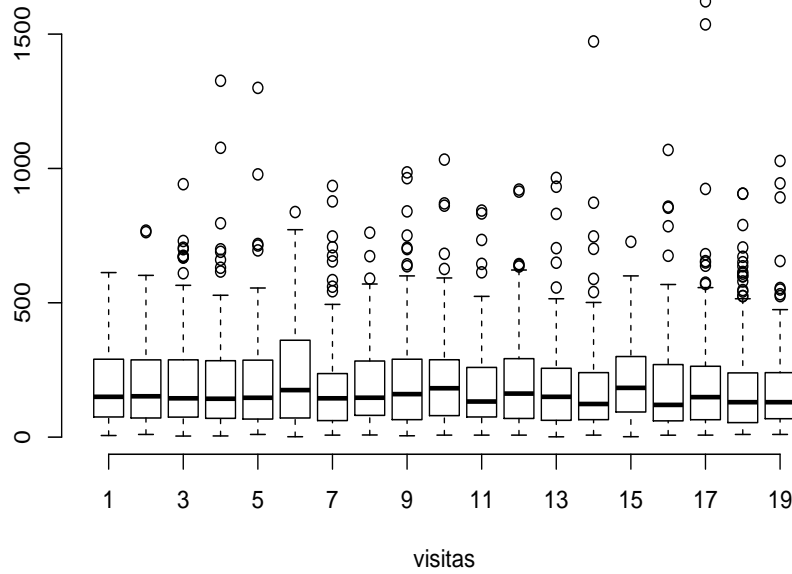


Figura 4.1: Distribuição do consumo positivo de FLV ao longo das visitas.

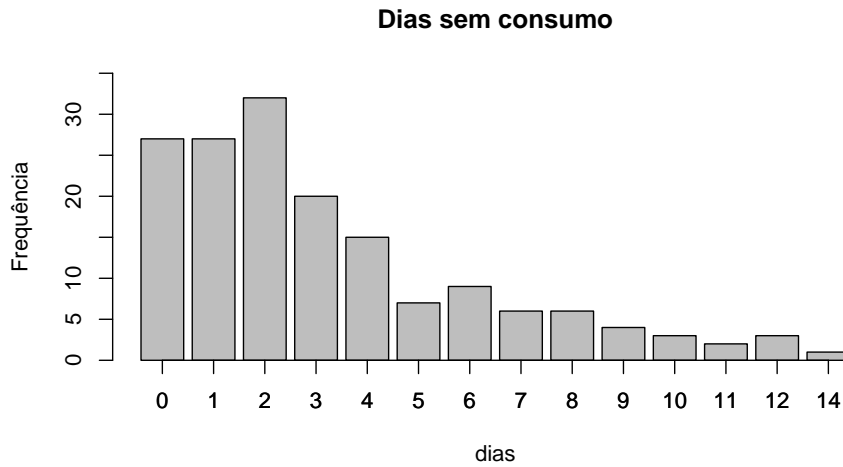


Figura 4.2: Frequência de dias sem consumo de FLV para os indivíduos do estudo.

onde $\mathbf{x}_{ij,\mu}^\top$ e $\mathbf{x}_{ij,p}^\top$ são matrizes de efeitos fixos associados ao indivíduo i que contém o intercepto e 8 covariáveis (idade, IMC, sexo, refrigerante, FLV adequado, renda, escolaridade e mudança de peso) denotadas por: $x_{i1} = 1$, x_{i2} (idade do indivíduo i), x_{i3} (IMC do indivíduo i), x_{i4} (0 se sexo masculino, 1 se sexo feminino), x_{i5} (0 se consome muito refrigerante, 1 se não), x_{i6} (0 se consome FLV adequadamente, 1 se não), x_{i7} (1 se renda entre 2.000-4.000, 0 caso contrário), x_{i8} (1 se renda entre 4.000-6.000, 0 caso contrário), x_{i9} (1 se renda > 6.0000 , 0 caso contrário), x_{i10}

(1 se possui Ensino Médio, 0 caso contrário), x_{i11} (1 se possui Ensino Superior, 0 caso contrário), x_{i12} (1 se nível de escolaridade igual a Outro, 0 caso contrário), x_{i13} (1 se peso aumentou, 0 caso contrário) e x_{i14} (1 se peso diminuiu, 0 caso contrário). Deste modo, utilizamos como referência a primeira categoria no caso das covariáveis discretas. Os vetores β_μ e β_p são os vetores de parâmetros desconhecidos associados aos efeitos fixos para a média e a probabilidade de zero, respectivamente. As quantidades $u_{1,i}$ e $u_{2,i}$ são os efeitos aleatórios de cada indivíduo associados à média e a probabilidade de zero, respectivamente, tal que $(u_{1,i}, u_{2,i}) \sim N_2(\mathbf{0}, \Sigma_u)$, com

$$\Sigma_u = \begin{pmatrix} \sigma_{u_1}^2 & \rho\sigma_{u_1}\sigma_{u_2} \\ \rho\sigma_{u_1}\sigma_{u_2} & \sigma_{u_2}^2 \end{pmatrix}. \quad (4.2.2)$$

A quantidade b_{ij} é o efeito aleatório associado ao indivíduo i e a observação j tal que $b_{ij} \sim N(0, \sigma_b^2)$. Aqui, b_{ij} é independente de $(u_{1,i}, u_{2,i})$.

Utilizando os resultados do estudo realizado na Seção 3.2, consideramos $\beta_{k,\mu} \sim N(0, 10^3)$ e $\beta_{k,p} \sim N(0, 10^3)$, $k = 1, \dots, 14$. Para a priori do parâmetro de dispersão tomamos $\phi \sim \text{Gama}(0, 01, 0, 01)$. Para os parâmetros de dispersão dos efeitos aleatórios $u_{1,i}$ e $u_{2,i}$, temos

$$u_{1,i} \sim N(0, \eta_1^2) \quad \text{e} \quad u_{2,i}|u_{1,i} \sim N(\psi u_{1,i}, \eta_2^2), \quad (4.2.3)$$

tal que

$$\Sigma_u = \begin{pmatrix} \eta_1^2 & \psi\eta_1^2 \\ \psi\eta_1^2 & \psi^2\eta_1^2 + \eta_2^2 \end{pmatrix}, \quad (4.2.4)$$

com $\psi \sim N(0, 10^3)$ e $\eta_1^2, \eta_2^2 \sim \text{Inv-Gama}(0, 01, 0, 01)$. Além disso, $\sigma_b^2 \sim \text{Inv-Gama}(0, 01, 0, 01)$.

Para o ajuste de cada modelo, foram usadas duas cadeias MCMC independentes de tamanho 100.000 cada começando em valores distantes, com aquecimento inicial de 50.000 pontos e espaçamento de tamanho 100, resultando em 500 observações da posteriori cada. Para melhorar a convergência, dividimos a resposta por 100 e as variáveis idade e IMC foram padronizadas.

Apresentamos os intervalos de credibilidade de 95% simétricos amostrais para os modelos utilizando as distribuições Gama, Inversa Gaussiana e Lognormal nas Figuras 4.3, 4.4 e 4.5. Consideramos como significativos no modelo os parâmetros em que o intervalo não contém o valor zero, apresentados em cinza nas Figuras 4.3, 4.4 e 4.5. Poderíamos utilizar aqui critérios

de comparação como o CPO e o DIC apresentados na Seção 2.3, porém, utilizamos intervalos de credibilidade para evitar reajustar o modelo ao incluir ou retirar um parâmetro no modelo como é necessário para o cálculo desses critérios.

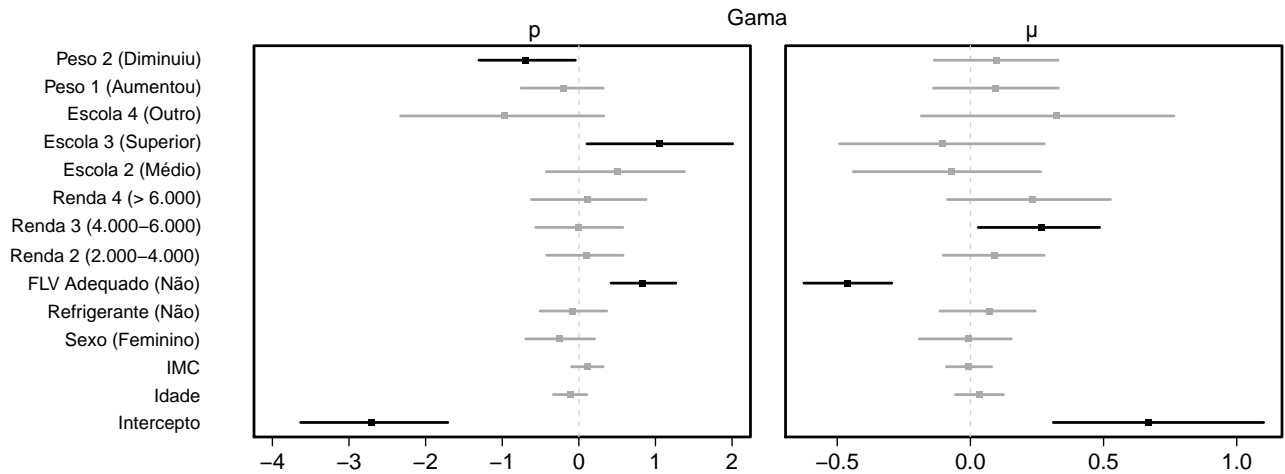


Figura 4.3: Intervalos de crebilidade de 95% simétricos para os parâmetros dos efeitos fixos para a probabilidade de valor zero (esquerda) e para a média dos valores positivos (direita) no modelo com a distribuição Gama.

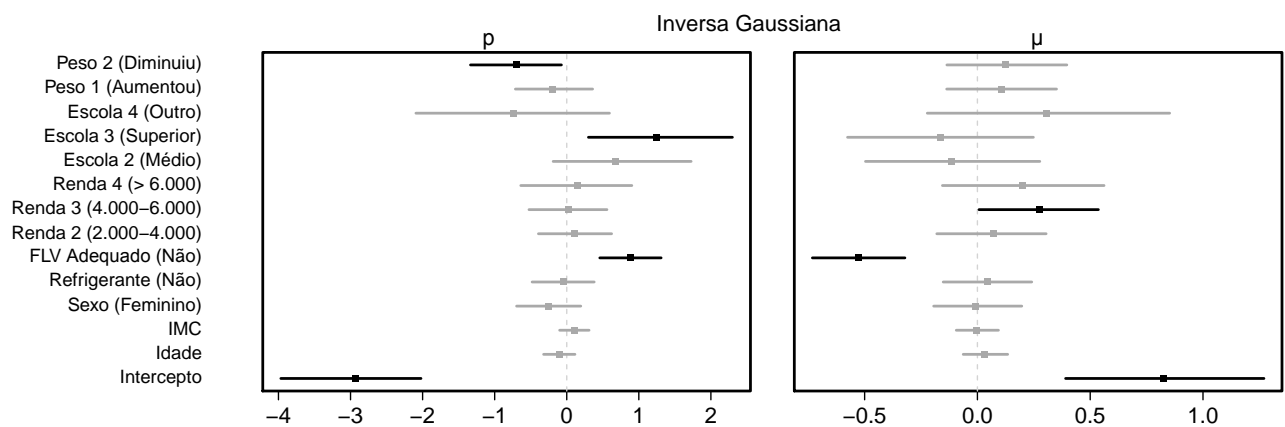


Figura 4.4: Intervalos de crebilidade de 95% simétricos para os parâmetros dos efeitos fixos para a probabilidade de valor zero (esquerda) e para a média dos valores positivos (direita) no modelo com a distribuição Inversa Gaussiana.

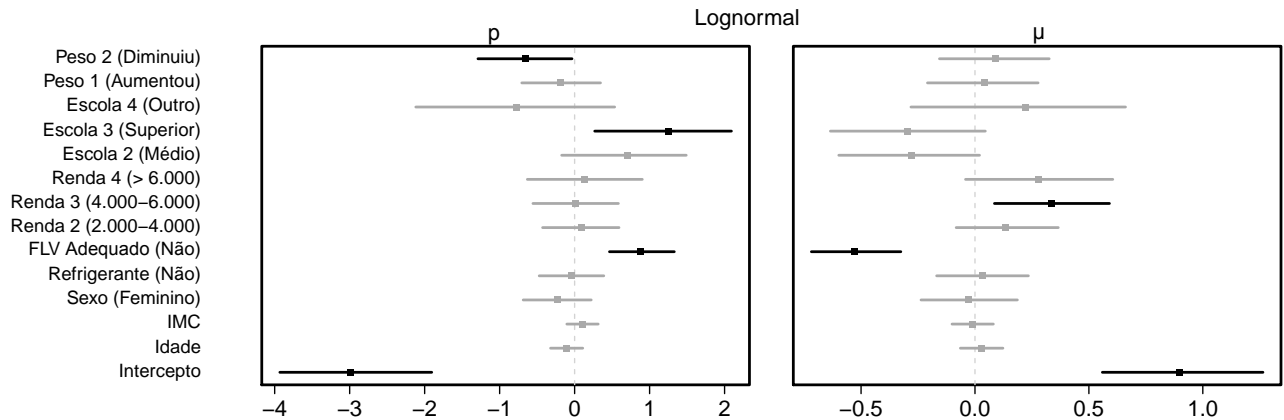


Figura 4.5: Intervalos de credibilidade de 95% simétricos para os parâmetros dos efeitos fixos para a probabilidade de valor zero (esquerda) e para a média dos valores positivos (direita) no modelo com a distribuição Lognormal.

Para os três modelos, temos que são significativas para a probabilidade de zero as variáveis FLV adequado, escola e peso e para a média do consumo positivo são significativas as variáveis FLV adequado e renda. No caso em que a covariável tem mais de uma categoria e apenas uma delas é significativa, mantemos no modelo todas as categorias. Assim, chegamos ao modelo final mantendo apenas as variáveis que foram significativas.

A Tabela 4.3 apresenta as estimativas da média, desvio-padrão (DP) e intervalos de credibilidade de 95% amostral dos parâmetros do modelo final para as distribuições Gama, Inversa Gaussiana e Lognormal. Podemos notar que as estimativas para os três modelos são bem próximas, com a maior diferença para o parâmetro ϕ , que tem média a posteriori mais alta para o modelo Gama. Para todos os ajustes, o intervalo de credibilidade para o parâmetro ρ não inclui zero, indicando correlação entre os efeitos aleatórios $u_{1,i}$ e $u_{2,i}$. O intervalo de credibilidade para o desvio-padrão σ_b de b_{ij} têm valor inferior próximo de zero, indicando que esse parâmetro pode ser próximo de zero.

A Tabela 4.3 apresenta também os critérios de comparação DIC_7 e LPML estudados na Seção 2.3 para os três modelos. Seguindo Congdon (2005), se as log-verossimilhanças utilizadas no cálculo dos CPOs excedem -10 , há uma ausência de indicação de problemas de instabilidade do estimador por média harmônica. Para cada observação, calculamos a porcentagem de vezes em que obtemos log-verossimilhanças superiores a -10 ao longo das iterações. O número de observações com log-verossimilhanças inferiores a -10 corresponde a 4, 7 e 2 para os mo-

delos Gama, Inversa Gaussiana e Lognormal, respectivamente. Esses resultados indicam que os estimadores para os CPOs não devem ter problemas de instabilidade e, portanto, podem ser utilizados na comparação dos modelos. Assim, de acordo com os critérios DIC_7 e LPML, o modelo com melhor ajuste é o modelo Gama, pois este possui o maior valor de LPML e o menor valor para o DIC_7 , como mostrado na Tabela 4.3. Dado esses critérios, selecionamos o modelo Gama como o melhor modelo.

Tabela 4.3: Estimativas a posteriori dos parâmetros e critérios de comparação para os modelos finais utilizando as distribuições Gama, Inversa Gaussiana e Lognormal.

Parâm. (Var.)	Gama				Inversa Gaussiana				Lognormal			
	Média	DP	2,5%	97,5%	Média	DP	2,5%	97,5%	Média	DP	2,5%	97,5%
Probabilidade de zero												
β_0 (Intercepto)	-2,88	0,39	-3,68	-2,16	-2,93	0,39	-3,74	-2,16	-2,84	0,40	-3,59	-2,06
β_1 (FLV Adequado - Não)	0,82	0,22	0,40	1,24	0,82	0,21	0,41	1,27	0,84	0,21	0,43	1,26
β_2 (Ensino Médio)	0,46	0,37	-0,27	1,22	0,54	0,38	-0,17	1,37	0,42	0,38	-0,29	1,17
β_3 (Ensino Superior)	0,96	0,39	0,22	1,69	1,03	0,39	0,28	1,87	0,93	0,40	0,17	1,75
β_4 (Ensino - Outro)	-0,64	0,56	-1,79	0,37	-0,65	0,56	-1,76	0,41	-0,66	0,58	-1,87	0,47
β_5 (Peso Aumentou)	-0,02	0,23	-0,45	0,39	0,01	0,23	-0,43	0,45	-0,07	0,23	-0,50	0,43
β_6 (Peso Diminuiu)	-0,63	0,28	-1,18	-0,08	-0,63	0,29	-1,19	-0,07	-0,61	0,29	-1,15	-0,10
Média												
α_0 (Intercepto)	0,69	0,10	0,50	0,88	0,75	0,12	0,52	0,98	0,70	0,11	0,48	0,91
α_1 (FLV Adequado - Não)	-0,46	0,08	-0,62	-0,30	-0,50	0,10	-0,70	-0,31	-0,50	0,09	-0,68	-0,33
α_2 (Renda 2.000-4.000)	0,12	0,09	-0,07	0,28	0,12	0,11	-0,09	0,32	0,17	0,10	-0,04	0,36
α_3 (Renda 4.000-6.000)	0,25	0,10	0,05	0,45	0,27	0,12	0,04	0,50	0,33	0,11	0,12	0,53
α_4 (Renda >6.000)	0,29	0,14	0,02	0,54	0,28	0,16	-0,04	0,58	0,36	0,14	0,08	0,63
Componentes de Variância												
σ_b	0,13	0,04	0,07	0,22	0,14	0,02	0,10	0,19	0,19	0,09	0,07	0,41
σ_{u_1}	0,45	0,03	0,39	0,51	0,46	0,04	0,38	0,54	0,48	0,03	0,42	0,55
σ_{u_2}	0,94	0,09	0,77	1,13	0,93	0,09	0,76	1,12	0,93	0,09	0,77	1,11
ρ	-0,64	0,09	-0,79	-0,46	-0,62	0,10	-0,79	-0,42	-0,61	0,09	-0,76	-0,42
ϕ	1,79	0,06	1,69	1,91	1,19	0,04	1,12	1,27	0,97	0,10	0,84	1,26
DIC_7	10276,42				11429,50				10482,75			
LPML	-5143,62				-5703,44				-5250,83			

Para analisar a convergência das cadeias nos três modelos, os valores da estatística \hat{R} são apresentados na Tabela C.1 no Apêndice C para o modelo final. Uma vez que nenhum dos valores ultrapassa 1,2, assumimos a convergência das cadeias. Também, o Apêndice C apresenta gráficos do tipo Trace plot nas Figuras C.1, C.2 e C.3 para os parâmetros do modelo final utilizando apenas a distribuição Gama, por simplicidade. Não há indícios de que a convergência

não foi atingida em nenhum dos casos.

As Figuras 4.6, 4.7 e 4.8 apresentam as medidas de divergência de Kullback-Liebler, a J-distância e a L_1 -distância para os modelos ajustados, utilizando como pontos de corte 0,51, 0,88 e 0,80, respectivamente. Podemos notar que temos medidas influentes para os três modelos, sendo o modelo que utiliza a distribuição Inversa Gaussiana o que tem o menor número de observações que ultrapassam os pontos de corte de todas as medidas de divergência. Devido aos ajustes dos modelos serem dispendiosos em tempo computacional e dada a quantidade de observações consideradas influentes em cada caso, não conduzimos o ajuste dos modelos retirando cada possível observação influente, mas apenas retirando todas as observações assinaladas de uma vez considerando o modelo utilizando a distribuição Gama, pois este foi escolhido como o melhor modelo. Observamos que, de modo geral, as estimativas e respectivo desvio padrão dos parâmetros não tiveram muita alteração, sendo mantidas as conclusões sobre a significância dos parâmetros. Assim, tomamos que as estimativas apresentadas para o modelo utilizando a distribuição Gama podem ser consideradas adequadas e continuamos a análise com todas as observações. Também, como uma abordagem à qualidade do ajuste, realizamos uma análise da capacidade preditiva do modelo escolhido como o melhor segundo os critérios de comparação utilizados.

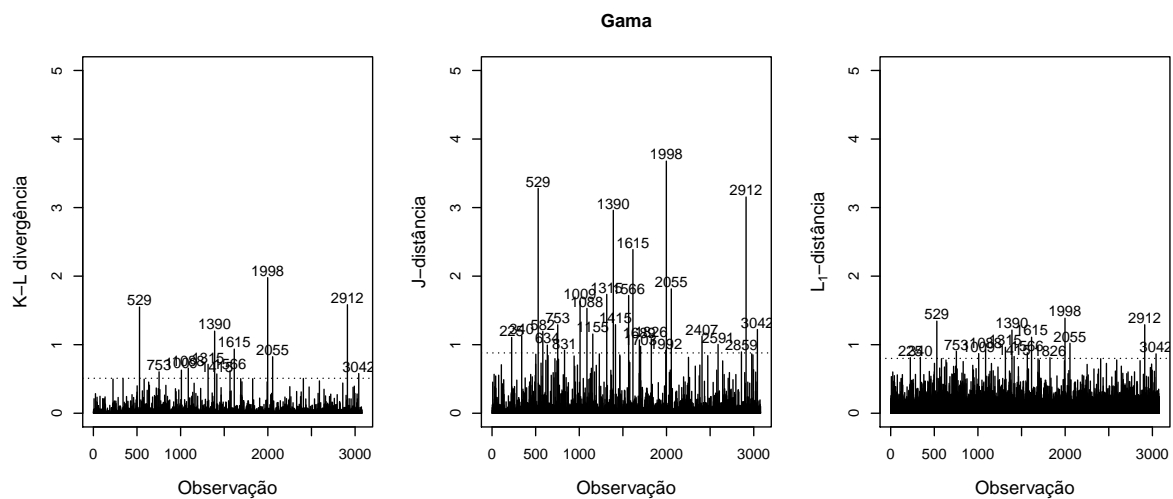


Figura 4.6: Medidas de divergência de Kullback-Liebler, J-distância e L_1 -distância para o modelo Gama.

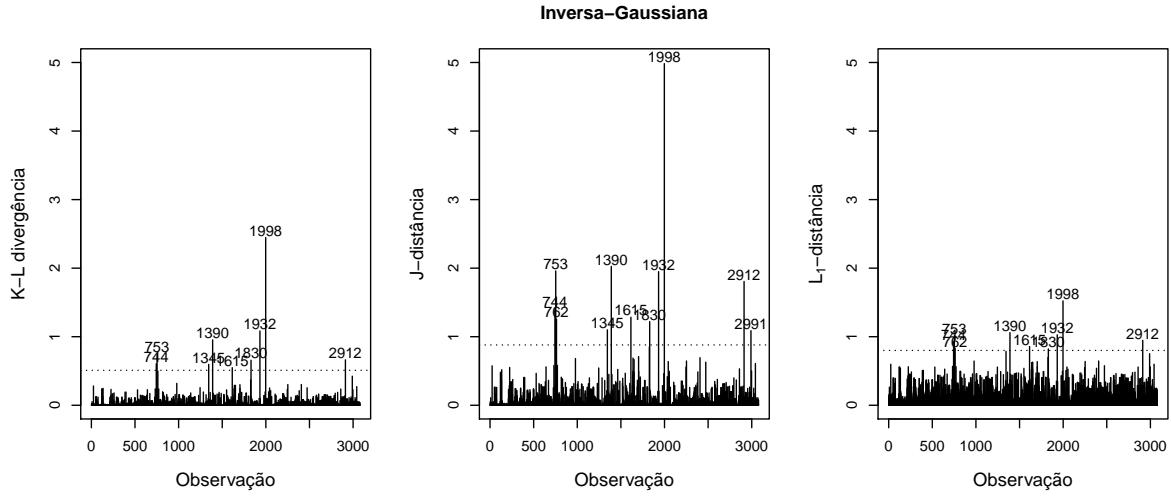


Figura 4.7: Medidas de divergência de Kullback-Liebler, J-distância e L_1 -distância para o modelo Inversa Gaussiana.

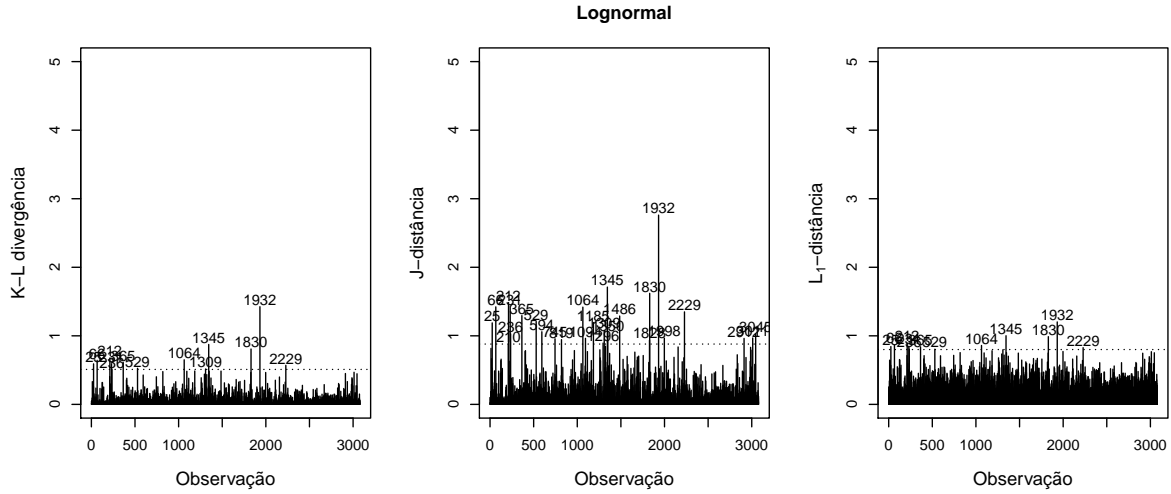


Figura 4.8: Medidas de divergência de Kullback-Liebler, J-distância e L_1 -distância para o modelo Lognormal.

Uma vez que escolhemos Gama como o melhor modelo, procedemos à interpretação dos parâmetros dos efeitos fixos do modelo Gama final. Tomando os efeitos aleatórios iguais a zero, temos

$$\mu_{ij} = \exp(\mathbf{x}_{ij}^\top \boldsymbol{\beta}_\mu) \quad \text{e} \quad p_{ij} = \exp\{\mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p\} / (1 + \exp\{\mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p\}). \quad (4.2.5)$$

Para a média dos valores positivos, lembrando que dividimos a resposta por 100 no ajuste, o parâmetro α_0 é o logaritmo da média (dividida por 100) para os indivíduos que conso-

mem FLV adequadamente e têm renda < 2.000 reais, ou seja, a média para esses indivíduos é igual a $\exp(\alpha_0) \times 100 = \exp(0,69) \times 100 = 199,37$ gramas. O parâmetro α_1 é o incremento (negativo) no logaritmo da média para os indivíduos que não consomem FLV adequadamente em relação aos que consomem e tem renda < 2.000 reais. Os parâmetros α_2, α_3 e α_4 representam o incremento (positivo) no logaritmo da média por ter renda nas categorias 2 (2.000-4.000), 3 (4.000-6.000) e 4 (> 6.000), respectivamente, em relação aos que tem renda < 2.000 reais e consomem FLV adequadamente. Por exemplo, a média estimada para um indivíduo que não consome FLV adequadamente e possui renda (> 6.000) reais é igual a $\exp(\alpha_0 + \alpha_1 + \alpha_4) \times 100 = \exp(0,69 - 0,46 + 0,29) \times 100 = 168,20$ gramas.

Para a probabilidade de zeros, $\text{logito}(p) = \log(p/(1-p))$ pode ser visto como o logaritmo da chance entre obter valor zero e obter valor positivo. Assim, β_0 é o logaritmo da chance entre obter valor zero e obter valor positivo para os indivíduos que consomem FLV adequadamente, possuem Ensino Fundamental e não mudaram de peso antes do estudo, ou seja, a probabilidade estimada de obter valor zero para esse grupo é igual a $\exp(\beta_0)/(1 + \exp(\beta_0)) = \exp(-2,88)/(1 + \exp(-2,88)) = 0,05$. Os parâmetros β_1, \dots, β_6 representam o incremento (positivo ou negativo) no logaritmo da chance entre obter valor zero e obter valor positivo em relação aos indivíduos que consomem FLV adequadamente, possuem Ensino Fundamental e não mudaram de peso antes do estudo. Por exemplo, para um indivíduo que possui Ensino Superior e aumentou de peso, a probabilidade de valor zero é dada por $\exp(\beta_0 + \beta_3 + \beta_5)/(1 + \exp(\beta_0 + \beta_3 + \beta_5)) = \exp(-2,88 + 0,96 - 0,02)/(1 + \exp(-2,88 + 0,96 - 0,02)) = 0,13$.

Para a média do consumo considerando os valores zero, podemos calcular sua estimativa considerando a média do consumo positivo e a probabilidade de zeros como dado em (1.1.39). Por exemplo, para os indivíduos que consomem FLV adequadamente, possuem Ensino Fundamental, não mudaram de peso antes do estudo e tem renda < 2.000 reais, o consumo médio, considerando apenas os efeitos fixos, é dado por $(1/(1 + \exp(\beta_0))) \exp(\alpha_0) \times 100 = (1/(1 + \exp(-2,88))) \exp(0,69) \times 100 = 188,78$ gramas.

Para avaliar o ajuste do modelo escolhido em termos da sua capacidade preditiva, geramos 1000 replicações da resposta $\mathbf{Y}^* = (\mathbf{Y}^{(1)}, \dots, \mathbf{Y}^{(1000)})^\top$, onde o elemento ij da l -ésima replicação $Y_{ij}^{(l)}$ é gerado através do modelo Gama aumentado em zero dado por $g(y_{ij}|\mu_{ij}^{(l)}, \phi^{(l)}, p_{ij}^{(l)})$, onde $\mu_{ij}^{(l)} = \exp(\mathbf{x}_{ij,\mu}^\top \boldsymbol{\beta}_\mu^{(l)} + u_{1,i}^{(l)} + b_{ij}^{(l)})$ e $p_{ij}^{(l)} = \exp\{\mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p^{(l)} + u_{2,i}^{(l)}\} / (1 + \exp\{\mathbf{x}_{ij,p}^\top \boldsymbol{\beta}_p^{(l)} + u_{2,i}^{(l)}\})$ para $\boldsymbol{\beta}_\mu^{(l)} = (\alpha_0^{(l)}, \dots, \alpha_4^{(l)})$ e $\boldsymbol{\beta}_p^{(l)} = (\beta_0^{(l)}, \dots, \beta_6^{(l)})$ e $\phi^{(l)}$ são as amostras de tamanho 1000 obtidas da distribuição a posteriori dos parâmetros, considerando o período de aquecimento e espaçamento. Para os valores positivos, a Figura 4.9 apresenta o histograma da densidade preditiva do consumo de FLV em todas as visitas e a curva da densidade preditiva a posteriori para o

modelo Gama, que sugere que o modelo produz um ajuste adequado aos dados. Além disso, apresentamos as curvas para os modelos utilizando a distribuição Inversa Gaussiana e Lognormal, que sugerem que o modelo Gama é o mais adequado.

Para os valores zero, calculamos o p-valor Bayesiano a partir da preditiva usando como medida de discrepância a proporção de zeros observados na amostra gerada em cada replicação, como sugerido em Neelon et al. (2010) para modelos zero inflacionados. Sendo a estimativa por Monte Carlo do p-valor Bayesiano dada pela proporção de vezes em que a proporção de zeros da amostra gerada foi maior que a a proporção de zeros observada, obtemos p-valor igual a 0,47 para as 1000 amostras geradas. Como esse valor é próximo de 0,5, temos evidências de um bom ajuste.

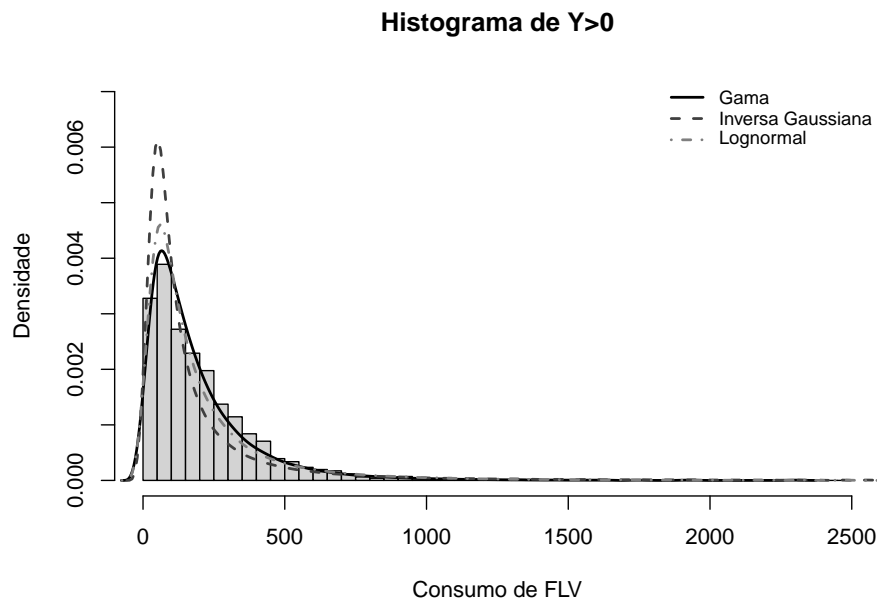


Figura 4.9: Histograma da densidade preditiva do Consumo positivo de FLV com as curvas para a densidade preditiva a posteriori gerada dos modelos Gama, Inversa Gaussiana e Lognormal sobrepostas.

Por fim, para verificar a adequação da função logaritmo usada como função link para modelar a média condicional dos valores positivos μ , Hatfield et al. (2012) sugere comparar as médias preditas a posteriori (box-plot em preto) e os valores positivos observados (box-plot em cinza) divididas em 10 intervalos com o número de observações parecido como apresentado na Figura 4.10. Uma vez que os box-plots das médias ajustadas e dos valores observados possuem um formato de tendência parecida a função link utilizada parece adequada.

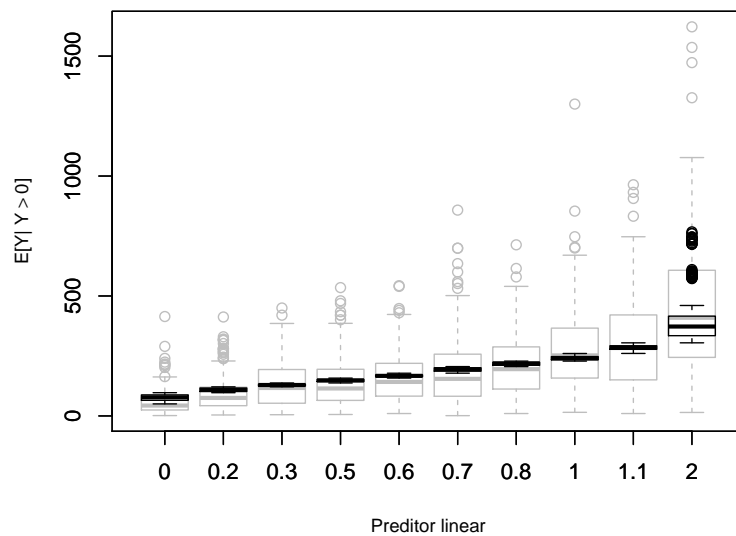


Figura 4.10: Box-plot das médias preditas a posteriori (em preto) e dos valores positivos observados (em cinza).

Capítulo 5

Considerações finais

Neste trabalho pudemos estudar modelos mistos para repostas positivas aumentadas em zero por meio de uma revisão da teoria desenvolvida para esta classe de modelos e de uma aplicação em dados reais. A classe de modelos estudada possui a vantagem de conseguir modelar a resposta positiva e o zero simultaneamente e permitir que a distribuição da resposta positiva esteja numa classe de distribuições que acomode distribuições assimétricas parametrizadas pela média. Modelamos o logito da probabilidade do zero e o logaritmo da média da distribuição positiva através de efeitos fixos e aleatórios que acomodam medidas repetidas, sendo possível estimar a variabilidade intra e entre unidades experimentais. A análise dos dados de consumo habitual mostrou, de maneira ilustrativa, uma aplicação real desta classe de modelos. Todo o processo de estimação, escolha do melhor modelo e respectiva adequabilidade foram feitos a partir de uma perspectiva Bayesiana. Através de um estudo de simulação detalhado pudemos verificar qual seria uma priori mais apropriada para as componentes de variância dos efeitos aleatórios. Um outro estudo de simulação indicou que as estimativas dos parâmetros melhoram a partir de uma combinação entre o número de unidades experimentais no estudo e número de observações para cada unidade.

Em geral, na literatura encontramos referências que modelam o consumo habitual a partir de transformações, como por exemplo transformação Box-Cox. O modelo proposto neste trabalho apresenta a vantagem de modelar a resposta respeitando sua natureza. Em particular, para os dados utilizados na aplicação, utilizamos o modelo de repostas positivas aumentadas em zero a partir das distribuições Gama, Inversa Gaussiana e Lognormal. Na aplicação os resultados foram muito próximos para a maioria dos parâmetros, exceto pelo parâmetro de dispersão da distribuição da resposta positiva. A partir dos critérios de seleção de modelos, o modelo a partir da distribuição Gama foi apontado como o melhor. No entanto, a partir de uma análise de pontos de influência notamos que o modelo com a distribuição Inversa Gaussiana é o que menos detectou pontos de influência, seguido pelos modelos com a distribuição

Lognormal e Gama. Dado este resultado, foi feita uma análise para verificar a capacidade preditiva do modelo a partir da distribuição Gama e verificou-se que este modelo produz um ajuste adequado aos dados. Finalmente foi verificada a adequação da função de ligação usadas no modelo para a média a partir da distribuição Gama e novamente o resultado foi favorável.

Uma sugestão para trabalhos futuros seria modelar a resposta através de distribuições que não pertencem à classe da família exponencial de dois parâmetros, como por exemplo definir o modelo de respostas positivas aumentadas em zero a partir da distribuição Birbaum-Saunders (Leiva et al., 2014). Outra contribuição seria permitir outras distribuições para os efeitos aleatórios. A presença de efeitos aleatórios implica em um processo intenso de predição, tal que uma maneira de contornar o problema seria modelar o logito da probabilidade do zero em função do logaritmo da média. Se o interesse maior é na estimativa das componentes de variância dos efeitos aleatórios e não na sua predição, uma outra proposta seria considerar distribuições a priori diretamente para o logaritmo da média e para o logito da probabilidade de se observar um valor zero. Finalmente, outra possível extensão seria não tomar o parâmetro de precisão ϕ como constante entre as observações, mas considerar a heterocedasticidade de ϕ e modelar este como função das covariáveis.

Referências

- [1] Aitchison, John (1955). “On the Distribution of a Positive Random Variable Having a Discrete Probability Mass at the Origin*”. Em: *Journal of the american statistical association* 50.271, pp. 901–908.
- [2] Bar-Lev, Shaul K e Benjamin Reiser (1982). “An exponential subfamily which admits UMPU tests based on a single test statistic”. Em: *The Annals of Statistics*, pp. 979–989.
- [3] Bose, Arup e Benzion Boukai (1993). “Sequential estimation results for a two-parameter exponential family of distributions”. Em: *The Annals of Statistics*, pp. 484–502.
- [4] Brown, Lawrence D (1986). “Fundamentals of statistical exponential families with applications in statistical decision theory”. Em: *Lecture Notes-monograph series*, pp. i–279.
- [5] Browne, William J e David Draper (2000). “Implementation and performance issues in the Bayesian and likelihood fitting of multilevel models”. Em: *Computational statistics* 15, pp. 391–420.
- [6] Carlin, Bradley P e Thomas A Louis (2008). *Bayesian methods for data analysis*. CRC Press.
- [7] Carriquiry, Alicia L (2003). “Estimation of usual intake distributions of nutrients and foods”. Em: *The Journal of nutrition* 133.2, 601S–608S.
- [8] Celeux, Gilles et al. (2006). “Deviance information criteria for missing data models”. Em: *Bayesian analysis* 1.4, pp. 651–673.
- [9] Chung, Yeojin et al. (2015). “Weakly informative prior for point estimation of covariance matrices in hierarchical models”. Em: *Journal of Educational and Behavioral Statistics* 40.2, pp. 136–157.
- [10] Congdon, Peter (2005). *Bayesian models for categorical data*. John Wiley & Sons.
- [11] Cook, R Dennis e Sanford Weisberg (1982). “Residuals and influence in regression”. Em:
- [12] Cowles, Mary Kathryn e Bradley P Carlin (1996). “Markov chain Monte Carlo convergence diagnostics: a comparative review”. Em: *Journal of the American Statistical Association* 91.434, pp. 883–904.
- [13] Dey, Dipak K, Alan E Gelfand e Fengchun Peng (1997). “Overdispersed generalized linear models”. Em: *Journal of Statistical Planning and Inference* 64.1, pp. 93–107.

- [14] Dodd, Kevin W et al. (2006). “Statistical methods for estimating usual intake of nutrients and foods: a review of the theory”. Em: *Journal of the American Dietetic Association* 106.10, pp. 1640–1650.
- [15] Duan, Naihua et al. (1983). “A comparison of alternative models for the demand for medical care”. Em: *Journal of business & economic statistics* 1.2, pp. 115–126.
- [16] Ferrari, Silvia e Francisco Cribari-Neto (2004). “Beta regression for modelling rates and proportions”. Em: *Journal of Applied Statistics* 31.7, pp. 799–815.
- [17] Galvis, Diana M, Dipankar Bandyopadhyay e Victor H Lachos (2014). “Augmented mixed beta regression models for periodontal proportion data”. Em: *Statistics in medicine* 33.21, pp. 3759–3771.
- [18] Geisser, Seymour (1993). *Predictive inference*. Vol. 55. CRC Press.
- [19] Geisser, Seymour e William F Eddy (1979). “A predictive approach to model selection”. Em: *Journal of the American Statistical Association* 74.365, pp. 153–160.
- [20] Gelfand, Alan E (1996). “Model determination using sampling-based methods”. Em: *Markov chain Monte Carlo in practice*, pp. 145–161.
- [21] Gelfand, Alan E e Adrian FM Smith (1990). “Sampling-based approaches to calculating marginal densities”. Em: *Journal of the American statistical association* 85.410, pp. 398–409.
- [22] Gelman, Andrew et al. (2006). “Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper)”. Em: *Bayesian analysis* 1.3, pp. 515–534.
- [23] Geman, Stuart e Donald Geman (1984). “Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images”. Em: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 6, pp. 721–741.
- [24] Givens, Geof H e Jennifer A Hoeting (2012). *Computational statistics*. Vol. 710. John Wiley & Sons.
- [25] Hatfield, Laura A et al. (2012). “Multilevel Bayesian models for survival times and longitudinal patient-reported outcomes with many zeros”. Em: *Journal of the American Statistical Association* 107.499, pp. 875–885.
- [26] Kipnis, Victor et al. (2009). “Modeling data with excess zeros and measurement error: application to evaluating relationships between episodically consumed foods and health outcomes”. Em: *Biometrics* 65.4, pp. 1003–1010.
- [27] Lambert, Diane (1992). “Zero-inflated Poisson regression, with an application to defects in manufacturing”. Em: *Technometrics* 34.1, pp. 1–14.
- [28] Lambert, Paul C et al. (2005). “How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS”. Em: *Statistics in medicine* 24.15, p. 2401.

- [29] Lee, Andy H et al. (2010). “How to analyze longitudinal multilevel physical activity data with many zeros?” Em: *Preventive medicine* 51.6, pp. 476–481.
- [30] Leiva, Víctor et al. (2014). “Birnbau-Saunders statistical modelling: a new approach”. Em: *Statistical Modelling* 14.1, pp. 21–48.
- [31] Lewis, Paul O et al. (2014). “Posterior predictive Bayesian phylogenetic model selection”. Em: *Systematic biology* 63.3, pp. 309–321.
- [32] Ligges, U et al. (2009). “BRugs 0.5: OpenBUGS and its R/S-PLUS interface BRugs”. Em: URL <http://www.stats.ox.ac.uk/pub/RWin/src/contrib>.
- [33] Liu, Lei, Jennie Z Ma e Bankole A Johnson (2008). “A multi-level two-part random effects model, with application to an alcohol-dependence study”. Em: *Statistics in medicine* 27.18, pp. 3528–3539.
- [34] McCulloch Charles, E e Shayle R Searle (2001). *Generalized, Linear, and Mixed Models*.
- [35] Neelon, Brian H, A James O’Malley e Sharon-Lise T Normand (2010). “A Bayesian model for repeated measures zero-inflated count data with application to outpatient psychiatric service use”. Em: *Statistical Modelling* 10.4, pp. 421–439.
- [36] Ntzoufras, Ioannis (2011). *Bayesian modeling using WinBUGS*. Vol. 698. John Wiley & Sons.
- [37] Olsen, Maren K e Joseph L Schafer (2001). “A two-part random-effects model for semicontinuous longitudinal data”. Em: *Journal of the American Statistical Association* 96.454, pp. 730–745.
- [38] Ospina, Raydonal e Silvia LP Ferrari (2012). “A general class of zero-or-one inflated beta regression models”. Em: *Computational Statistics & Data Analysis* 56.6, pp. 1609–1623.
- [39] Raftery, Adrian E et al. (2006). “Estimating the integrated likelihood via posterior simulation using the harmonic mean identity”. Em: *Bayesian Statistics (eds J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith and M. West))*, Vol. 8. Oxford University Press: Oxford, pp. 1–45.
- [40] Rodrigues-Motta, Mariana et al. (2015). “A mixed-effect model for positive responses augmented by zeros”. Em: *Statistics in medicine* 34.10, pp. 1761–1778.
- [41] Sinharay, Sandip e Hal S Stern (2005). “An empirical comparison of methods for computing Bayes factors in generalized linear mixed models”. Em: *Journal of Computational and Graphical Statistics* 14.2, pp. 415–435.
- [42] Spiegelhalter, David J et al. (2002). “Bayesian measures of model complexity and fit”. Em: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64.4, pp. 583–639.
- [43] Thomas, Andrew et al. (2006). “Making BUGS open”. Em: *R news* 6.1, pp. 12–17.

- [44] Tooze, Janet A, Gary K Grunwald e Richard H Jones (2002). “Analysis of repeated measures data with clumping at zero”. Em: *Statistical methods in medical research* 11.4, pp. 341–355.
- [45] Tooze, Janet A et al. (2006). “A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution”. Em: *Journal of the American Dietetic Association* 106.10, pp. 1575–1587.
- [46] Tooze, Janet A et al. (2010). “A mixed-effects model approach for estimating the distribution of usual intake of nutrients: The NCI method”. Em: *Statistics in medicine* 29.27, pp. 2857–2868.
- [47] Verly Junior, Eliseu et al. (2010). “Sources of variation of energy and nutrient intake among adolescents in São Paulo, Brazil”. Em: *Cadernos de Saúde Pública* 26.11, pp. 2129–2137.
- [48] Yau, Kelvin KW, Andy H Lee e Angus SK Ng (2002). “Theory & Methods: A zero-augmented gamma mixed model for longitudinal data with many zeros”. Em: *Australian & New Zealand Journal of Statistics* 44.2, pp. 177–183.
- [49] Zhang, Min et al. (2006). “Bayesian inference for a two-part hierarchical model: An application to profiling providers in managed health care”. Em: *Journal of the American Statistical Association* 101.475, pp. 934–945.

Apêndice A

Estudo de simulação para as prioris

Resultados da simulação para comparar as prioris dos tipos I, II, III e IV com respeito ao erro relativo (ER), erro quadrático médio (EQM) e probabilidade de cobertura (PC).

Tabela A.1: Resultados da simulação segundo ER, EQM e PC para $\rho = 0$.

Parâmetro	Priori					
	I, IV		II		III	
	$n = 50$	$n = 150$	$n = 50$	$n = 150$	$n = 50$	$n = 150$
ER						
β_0	2,04	0,16	4,45	0,26	3,87	0,57
β_1	0,11	-0,04	1,78	0,02	1,99	0,26
α_0	-0,02	-0,02	-0,02	-0,03	-0,05	-0,03
α_1	0,00	0,01	0,00	0,01	0,00	0,01
σ_b	0,19	0,01	0,27	0,07	0,79	0,40
σ_{u_1}	1,07	0,67	0,81	0,50	1,88	1,11
σ_{u_2}	5,86	1,27	19,68	2,50	19,72	5,14
ϕ	1,95	0,90	2,81	0,95	5,91	2,18
EQM						
β_0	135,97	0,44	367,64	0,76	161,82	2,29
β_1	245,71	0,38	316,55	0,43	231,43	0,85
α_0	0,03	0,02	0,03	0,03	0,04	0,02
α_1	0,03	0,02	0,03	0,02	0,03	0,01
σ_b	0,01	0,01	0,02	0,02	0,06	0,04
σ_{u_1}	0,01	0,01	0,01	0,02	0,04	0,03
σ_{u_2}	17,04	0,38	122,06	1,19	41,98	2,80
ϕ	83,69	42,43	160,14	40,86	389,63	144,34
PC						
β_0	0,91	0,92	0,75	0,87	0,58	0,76
β_1	0,89	0,96	0,91	0,97	0,90	0,94
α_0	0,96	0,94	0,96	0,91	0,96	0,91
α_1	0,94	0,91	0,95	0,90	0,96	0,93
σ_b	0,99	0,97	0,99	0,93	0,74	0,83
σ_{u_1}	0,95	0,96	0,98	1,00	0,50	0,78
σ_{u_2}	0,97	1,00	0,70	0,96	0,00	0,04
ϕ	0,93	0,90	0,94	0,88	0,62	0,76

Tabela A.2: Resultados da simulação segundo ER, EQM e PC para $\rho = 0,5$.

Parâmetro	Priori							
	I		II		III		IV	
	$n = 50$	$n = 150$	$n = 50$	$n = 150$	$n = 50$	$n = 150$	$n = 50$	$n = 150$
ER								
β_0	7,27	1,21	7,97	0,24	7,09	0,55	6,43	0,31
β_1	-4,35	1,00	-4,53	0,12	-4,10	0,33	-5,21	0,18
α_0	-0,01	-0,02	-0,01	-0,01	-0,04	-0,02	0,00	-0,01
α_1	-0,04	-0,01	-0,03	-0,01	-0,04	-0,01	-0,03	-0,01
σ_b	0,30	0,01	0,27	0,03	0,89	0,36	0,19	0,02
σ_{u_1}	15,29	9,22	0,95	0,31	2,56	1,00	0,84	0,40
σ_{u_2}	20,07	10,75	20,57	2,26	19,06	5,25	12,09	3,22
ρ	-1,01	-1,01	-1,04	-1,00	-1,02	-0,98	-1,05	-0,95
ϕ	2,73	1,40	2,71	0,91	6,48	2,02	1,87	0,85
EQM								
β_0	656,37	6,82	795,92	0,62	615,02	2,04	626,32	0,80
β_1	512,89	2,00	604,83	0,36	578,01	0,61	623,08	0,39
α_0	0,05	0,01	0,04	0,01	0,06	0,01	0,04	0,01
α_1	0,06	0,01	0,04	0,01	0,05	0,01	0,05	0,01
σ_b	0,01	0,01	0,02	0,02	0,13	0,02	0,02	0,01
σ_{u_1}	2,39	0,85	0,01	0,00	0,22	0,01	0,01	0,00
σ_{u_2}	42,57	10,54	136,57	0,76	41,07	2,89	30,86	1,17
ρ	0,25	0,25	0,32	0,29	0,32	0,31	0,40	0,39
ϕ	380,43	71,22	131,30	35,38	523,66	101,38	74,37	35,85
PC								
β_0	0,50	0,31	0,63	0,94	0,57	0,78	0,70	0,90
β_1	0,83	0,88	0,83	0,98	0,85	0,94	0,83	0,97
α_0	1,00	1,00	0,94	0,96	0,92	0,93	0,94	0,95
α_1	1,00	1,00	0,93	0,93	0,93	0,96	0,92	0,96
σ_b	1,00	1,00	0,99	0,91	0,74	0,84	0,99	0,97
σ_{u_1}	0,00	0,00	0,97	0,99	0,37	0,79	0,98	1,00
σ_{u_2}	0,00	0,00	0,58	0,85	0,01	0,10	0,41	0,66
ρ	0,01	0,00	0,99	1,00	0,95	0,92	0,99	1,00
ϕ	0,97	0,90	0,93	0,90	0,63	0,75	0,95	0,92

Tabela A.3: Resultados da simulação segundo ER, EQM e PC para $\rho = 0, 1$.

Parâmetro	Priori							
	I		II		III		IV	
	$n = 50$	$n = 150$	$n = 50$	$n = 150$	$n = 50$	$n = 150$	$n = 50$	$n = 150$
ER								
β_0	3,61	1,24	3,44	0,26	3,45	0,60	2,74	0,32
β_1	1,57	1,01	1,60	0,16	1,81	0,34	0,51	0,17
α_0	-0,02	-0,03	-0,03	-0,01	-0,05	-0,03	-0,01	-0,01
α_1	-0,01	0,01	-0,02	0,00	-0,01	0,01	-0,01	0,00
σ_b	0,27	-0,01	0,30	0,00	0,76	0,39	0,17	0,00
σ_{u_1}	15,23	9,22	1,37	0,49	2,03	1,26	0,85	0,50
σ_{u_2}	19,17	10,77	16,32	2,32	18,52	5,35	11,70	3,03
ρ	-0,99	-1,03	-0,92	-0,87	-0,65	-0,95	-0,53	-0,84
ϕ	2,09	1,11	2,28	0,83	6,48	2,06	1,88	0,70
EQM								
β_0	177,98	7,48	265,55	0,99	150,02	3,11	194,57	1,00
β_1	253,98	2,05	346,32	0,60	215,50	0,85	251,67	0,49
α_0	0,03	0,01	0,04	0,01	0,03	0,01	0,03	0,01
α_1	0,04	0,01	0,04	0,01	0,04	0,01	0,04	0,01
σ_b	0,01	0,00	0,04	0,01	0,06	0,04	0,01	0,01
σ_{u_1}	2,35	0,85	0,11	0,01	0,04	0,04	0,01	0,00
σ_{u_2}	35,89	10,65	92,70	0,97	37,59	3,25	42,20	1,09
ρ	0,01	0,01	0,05	0,05	0,06	0,07	0,14	0,15
ϕ	80,60	46,72	86,25	31,03	487,76	102,74	72,61	25,99
PC								
β_0	0,60	0,29	0,86	0,89	0,67	0,71	0,85	0,84
β_1	0,90	0,89	0,92	0,95	0,90	0,95	0,89	0,94
α_0	1,00	1,00	0,95	0,96	0,96	0,93	0,94	0,94
α_1	1,00	1,00	0,91	0,94	0,94	0,98	0,92	0,95
σ_b	1,00	1,00	0,99	0,95	0,77	0,94	1,00	0,98
σ_{u_1}	0,00	0,00	0,98	0,96	0,43	0,67	0,98	0,97
σ_{u_2}	0,00	0,00	0,56	0,87	0,00	0,08	0,40	0,78
ρ	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
ϕ	0,93	0,94	0,91	0,95	0,58	0,75	0,94	0,93

Tabela A.4: Resultados da simulação segundo ER, EQM e PC para $\rho = 0,9$.

Parâmetro	Priori							
	I		II		III		IV	
	$n = 50$	$n = 150$	$n = 50$	$n = 150$	$n = 50$	$n = 150$	$n = 50$	$n = 150$
ER								
β_0	6,84	1,39	7,03	0,48	6,49	0,69	5,94	0,46
β_1	-4,77	0,77	-4,76	-0,33	-4,11	0,11	-6,42	-0,08
α_0	-0,02	-0,03	-0,02	-0,02	-0,04	-0,04	-0,02	-0,02
α_1	0,00	0,01	-0,01	0,02	-0,01	0,02	-0,02	0,02
σ_b	0,32	0,02	0,24	0,00	0,80	0,34	0,21	-0,02
σ_{u_1}	15,09	9,28	0,78	0,35	2,06	1,04	0,84	0,44
σ_{u_2}	20,08	10,80	18,75	2,15	19,36	5,17	11,33	3,11
ρ	-1,00	-1,01	-0,99	-1,02	-0,98	-1,03	-0,97	-1,06
ϕ	2,62	1,29	2,76	0,80	5,74	1,78	1,78	0,65
EQM								
β_0	567,93	23,70	633,87	30,01	496,77	15,45	558,78	12,00
β_1	515,94	16,65	533,62	29,22	447,22	13,54	563,75	11,40
α_0	0,04	0,01	0,03	0,01	0,04	0,02	0,04	0,01
α_1	0,04	0,01	0,04	0,01	0,04	0,01	0,04	0,01
σ_b	0,01	0,01	0,02	0,02	0,10	0,02	0,02	0,01
σ_{u_1}	2,28	0,87	0,01	0,00	0,08	0,01	0,01	0,00
σ_{u_2}	40,09	10,66	107,86	0,65	40,60	2,77	20,61	1,06
ρ	0,81	0,82	0,83	0,90	0,84	0,92	0,90	1,04
ϕ	365,06	67,51	150,34	29,91	375,54	68,42	73,42	23,76
PC								
β_0	0,50	0,33	0,66	0,91	0,56	0,77	0,72	0,87
β_1	0,81	0,85	0,82	0,91	0,86	0,93	0,80	0,91
α_0	1,00	1,00	0,95	0,92	0,93	0,90	0,96	0,91
α_1	1,00	1,00	0,93	0,96	0,95	0,94	0,93	0,94
σ_b	1,00	0,99	0,97	0,93	0,78	0,92	0,98	0,95
σ_{u_1}	0,00	0,00	0,99	0,96	0,53	0,74	1,00	0,98
σ_{u_2}	0,00	0,00	0,60	0,81	0,01	0,11	0,42	0,69
ρ	0,00	0,00	0,57	0,32	0,28	0,19	0,97	0,93
ϕ	0,96	0,95	0,91	0,92	0,66	0,81	0,94	0,94

Tabela A.5: Resultados da simulação segundo ER, EQM e PC para $\rho = -0,5$.

Parâmetro	Priori							
	I		II		III		IV	
	$n = 50$	$n = 150$	$n = 50$	$n = 150$	$n = 50$	$n = 150$	$n = 50$	$n = 150$
ER								
β_0	5,41	1,20	4,76	0,23	5,36	0,51	3,92	0,31
β_1	-1,21	0,88	-2,12	0,07	-2,02	0,29	-1,87	0,13
α_0	-0,01	-0,02	-0,02	-0,01	-0,05	-0,03	-0,01	-0,01
α_1	-0,02	0,01	-0,02	0,01	-0,02	0,01	-0,02	0,01
σ_b	0,29	-0,03	0,33	-0,06	0,79	0,33	0,25	-0,05
σ_{u_1}	15,02	9,23	0,79	0,41	1,86	1,10	0,75	0,47
σ_{u_2}	18,35	10,41	11,24	2,11	17,58	4,94	8,39	3,04
ρ	-1,01	-0,97	-1,03	-1,04	-1,04	-1,05	-1,02	-1,08
ϕ	1,97	1,06	2,91	0,61	6,56	2,09	2,03	0,57
EQM								
β_0	419,79	7,19	422,16	0,62	404,96	1,82	329,92	0,85
β_1	551,54	2,04	589,65	0,48	431,38	0,80	541,22	0,55
α_0	0,03	0,01	0,03	0,01	0,04	0,01	0,03	0,01
α_1	0,04	0,01	0,04	0,01	0,04	0,01	0,04	0,01
σ_b	0,01	0,01	0,03	0,01	0,06	0,02	0,02	0,01
σ_{u_1}	2,26	0,92	0,01	0,00	0,04	0,01	0,01	0,00
σ_{u_2}	31,69	10,24	22,36	0,66	32,98	2,56	8,40	1,06
ρ	0,25	0,24	0,30	0,32	0,32	0,33	0,39	0,42
ϕ	54,42	43,62	143,30	18,68	504,56	117,54	76,99	25,17
PC								
β_0	0,59	0,22	0,79	0,95	0,64	0,77	0,78	0,91
β_1	0,83	0,86	0,86	0,97	0,85	0,95	0,86	0,95
α_0	1,00	0,97	0,97	0,92	0,96	0,96	0,96	0,94
α_1	1,00	0,97	0,94	0,97	0,97	0,98	0,94	0,97
σ_b	1,00	0,97	0,96	0,94	0,69	0,92	0,98	0,95
σ_{u_1}	0,00	0,00	0,99	0,96	0,48	0,74	1,00	0,99
σ_{u_2}	0,00	0,00	0,62	0,84	0,01	0,09	0,43	0,78
ρ	0,00	0,00	1,00	0,98	0,96	0,93	1,00	0,96
ϕ	0,96	0,90	0,91	0,93	0,49	0,88	0,94	0,95

Tabela A.6: Resultados da simulação segundo ER, EQM e PC para $\rho = -0,9$.

Parâmetro	Priori							
	I		II		III		IV	
	$n = 50$	$n = 150$	$n = 50$	$n = 150$	$n = 50$	$n = 150$	$n = 50$	$n = 150$
ER								
β_0	4,28	1,42	4,58	0,53	4,34	0,85	3,20	0,63
β_1	-0,09	0,64	-0,44	-0,42	-0,43	-0,20	-1,14	-0,41
α_0	-0,03	-0,02	-0,03	-0,01	-0,06	-0,02	-0,02	0,00
α_1	-0,01	0,00	0,00	0,00	0,00	0,00	-0,01	0,00
σ	0,34	-0,05	0,29	-0,06	0,85	0,30	0,22	-0,07
σ_{u_1}	15,29	9,22	0,73	0,40	2,19	1,04	0,77	0,44
σ_{u_2}	18,93	10,82	16,82	2,40	18,34	5,37	9,70	3,22
ρ	-0,99	-0,99	-0,99	-0,97	-0,97	-0,98	-0,94	-0,97
ϕ	2,34	0,93	2,78	0,55	6,43	1,65	1,94	0,31
EQM								
β_0	245,45	20,27	386,27	29,88	272,28	39,44	242,22	36,05
β_1	284,25	13,44	295,37	29,32	278,00	36,83	343,79	35,39
α_0	0,04	0,01	0,03	0,01	0,04	0,01	0,04	0,01
α_1	0,04	0,01	0,04	0,01	0,04	0,01	0,04	0,01
σ	0,02	0,00	0,02	0,01	0,10	0,02	0,02	0,01
σ_{u_1}	2,40	0,85	0,01	0,00	0,16	0,01	0,01	0,00
σ_{u_2}	34,77	10,72	82,52	0,78	36,32	2,98	13,73	1,17
ρ	0,80	0,80	0,84	0,82	0,81	0,84	0,84	0,90
ϕ	95,78	24,89	140,77	14,50	501,76	61,45	87,26	4,98
PC								
β_0	0,56	0,33	0,76	0,93	0,62	0,71	0,80	0,89
β_1	0,90	0,90	0,90	0,93	0,88	0,93	0,90	0,94
α_0	1,00	1,00	0,97	0,94	0,96	0,94	0,97	0,97
α_1	1,00	1,00	0,97	0,97	0,97	0,97	0,96	0,98
σ	1,00	1,00	0,99	0,96	0,83	0,96	1,00	0,96
σ_{u_1}	0,00	0,00	0,98	0,99	0,51	0,73	0,98	1,00
σ_{u_2}	0,00	0,00	0,55	0,81	0,00	0,06	0,39	0,69
ρ	0,00	0,00	0,60	0,38	0,31	0,20	0,96	0,93
ϕ	0,98	0,94	0,92	0,96	0,62	0,82	0,95	0,97

Apêndice B

Estudo de simulação para o número de observações por indivíduo

Resultados da simulação para comparar o erro relativo (ER), o erro quadrático médio (EQM) e a probabilidade de cobertura (PC) quando aumentamos o número de observações por indivíduo utilizando a priori do tipo IV.

Tabela B.1: Resultados da simulação segundo ER, EQM e PC para $\rho = 0,5$.

Parâmetro	Visitas = 5		Visitas = 10		Visitas = 15	
	$n = 50$	$n = 150$	$n = 50$	$n = 150$	$n = 50$	$n = 150$
ER						
β_0	0,38	0,09	0,08	0,05	0,07	0,02
β_1	-0,03	0,01	0,01	-0,01	0,02	0,01
α_0	-0,03	0,00	0,01	0,00	0,01	0,00
α_1	0,03	-0,01	-0,04	0,00	-0,03	0,00
σ_b	0,07	-0,07	-0,02	-0,10	-0,01	-0,01
σ_{u_1}	0,39	0,17	0,21	0,03	0,09	-0,03
σ_{u_2}	2,62	0,99	1,07	0,50	0,80	0,33
ρ	-0,90	-0,86	-0,85	-0,81	-0,65	-0,52
ϕ	0,79	0,09	0,22	-0,01	0,05	0,05
EQM						
β_0	1,52	0,14	0,17	0,04	0,11	0,03
β_1	0,88	0,14	0,20	0,05	0,12	0,04
α_0	0,01	0,00	0,01	0,00	0,01	0,00
α_1	0,01	0,00	0,01	0,00	0,01	0,00
σ_b	0,01	0,01	0,01	0,00	0,00	0,00
σ_{u_1}	0,00	0,00	0,00	0,00	0,00	0,00
σ_{u_2}	0,88	0,12	0,14	0,04	0,08	0,02
ρ	0,37	0,32	0,30	0,30	0,26	0,25
ϕ	21,63	2,65	2,63	0,08	0,30	0,24
PC						
β_0	0,90	0,94	0,92	0,96	0,96	1,00
β_1	0,92	0,94	0,94	0,94	0,96	0,98
α_0	0,94	0,96	0,98	1,00	0,92	0,92
α_1	0,98	1,00	0,96	0,96	0,90	0,96
σ_b	0,98	0,94	0,88	0,94	0,98	0,86
σ_{u_1}	0,98	1,00	1,00	0,96	1,00	0,98
σ_{u_2}	0,80	0,94	0,96	0,92	0,96	0,96
ρ	0,98	1,00	1,00	1,00	0,98	0,94
ϕ	0,98	0,94	0,94	0,90	0,96	0,92

Tabela B.2: Resultados da simulação segundo ER, EQM e PC para $\rho = 0, 1$.

	Visitas = 5		Visitas = 10		Visitas = 15	
	$n = 50$	$n = 150$	$n = 50$	$n = 150$	$n = 50$	$n = 150$
Parâmetro	ER					
β_0	0,20	0,10	0,07	0,04	0,07	0,01
β_1	-0,02	-0,05	0,11	0,02	0,05	0,04
α_0	-0,01	-0,01	0,00	0,00	0,01	0,00
α_1	-0,01	-0,01	-0,01	0,00	-0,01	0,00
σ_b	0,03	0,04	0,02	-0,08	-0,08	0,00
σ_{u_1}	0,32	0,22	0,17	0,00	0,10	0,00
σ_{u_2}	1,93	1,28	1,20	0,52	0,84	0,23
ρ	-1,06	-2,01	-2,51	-0,19	-0,41	-0,48
ϕ	1,11	0,11	0,49	0,00	0,04	0,19
Parâmetro	EQM					
β_0	0,52	0,13	0,18	0,04	0,08	0,02
β_1	0,48	0,14	0,24	0,06	0,11	0,04
α_0	0,01	0,03	0,01	0,00	0,01	0,00
α_1	0,01	0,01	0,01	0,00	0,01	0,00
σ_b	0,01	0,01	0,01	0,00	0,01	0,00
σ_{u_1}	0,00	0,00	0,00	0,00	0,00	0,00
σ_{u_2}	0,38	0,37	0,17	0,03	0,09	0,01
ρ	0,13	0,19	0,22	0,19	0,16	0,13
ϕ	55,53	0,73	17,23	0,09	0,44	6,44
Parâmetro	PC					
β_0	0,96	0,92	0,94	0,96	1,00	0,96
β_1	0,98	0,96	0,96	0,94	0,98	0,96
α_0	0,96	0,96	1,00	0,94	0,96	0,96
α_1	0,96	0,92	0,96	0,94	0,96	0,94
σ_b	1,00	0,96	0,90	0,92	0,92	0,96
σ_{u_1}	1,00	0,96	0,98	1,00	1,00	0,96
σ_{u_2}	0,96	0,98	0,94	0,96	0,92	1,00
ρ	1,00	0,98	0,98	1,00	0,98	1,00
ϕ	0,98	0,96	0,90	0,94	0,94	0,96

Tabela B.3: Resultados da simulação segundo ER, EQM e PC para $\rho = 0,9$.

Parâmetro	Visitas = 5		Visitas = 10		Visitas = 15	
	$n = 50$	$n = 150$	$n = 50$	$n = 150$	$n = 50$	$n = 150$
ER						
β_0	0,24	0,08	0,13	0,07	0,07	0,03
β_1	0,08	0,00	0,05	-0,09	-0,01	-0,04
α_0	-0,01	0,00	0,00	-0,01	0,00	0,00
α_1	-0,01	0,00	0,01	-0,01	0,00	0,00
σ_b	0,14	-0,11	-0,06	-0,05	-0,07	-0,02
σ_{u_1}	0,34	0,17	0,18	0,03	0,14	0,01
σ_{u_2}	2,39	1,00	1,20	0,74	0,81	0,31
ρ	-0,91	-0,82	-0,79	-0,63	-0,64	-0,51
ϕ	0,87	0,03	0,15	0,00	0,03	0,15
EQM						
β_0	0,72	0,16	0,26	0,08	0,16	0,03
β_1	0,45	0,17	0,24	0,07	0,13	0,04
α_0	0,01	0,01	0,01	0,02	0,00	0,00
α_1	0,02	0,01	0,01	0,01	0,00	0,00
σ_b	0,02	0,01	0,01	0,00	0,01	0,00
σ_{u_1}	0,00	0,00	0,00	0,00	0,00	0,00
σ_{u_2}	0,62	0,11	0,16	0,17	0,09	0,02
ρ	0,76	0,70	0,64	0,44	0,45	0,31
ϕ	27,69	0,28	1,65	0,15	0,18	4,13
PC						
β_0	0,92	0,90	0,92	0,90	0,98	0,96
β_1	0,96	0,88	0,94	0,98	0,98	0,94
α_0	0,98	0,96	0,94	0,98	0,98	1,00
α_1	0,94	0,94	0,98	0,96	0,94	1,00
σ_b	1,00	0,96	0,98	1,00	0,94	0,92
σ_{u_1}	1,00	1,00	1,00	0,98	0,98	0,96
σ_{u_2}	0,84	0,94	0,96	0,92	0,90	0,98
ρ	0,96	0,92	0,98	0,92	0,96	0,90
ϕ	0,98	0,94	0,96	0,98	0,94	0,96

Tabela B.4: Resultados da simulação segundo ER, EQM e PC para $\rho = -0,5$.

Parâmetro	Visitas = 5		Visitas = 10		Visitas = 15	
	$n = 50$	$n = 150$	$n = 50$	$n = 150$	$n = 50$	$n = 150$
ER						
β_0	0,33	0,10	0,05	0,03	0,05	0,00
β_1	0,24	0,02	0,07	-0,02	0,11	0,08
α_0	0,00	0,00	0,00	0,00	-0,01	-0,01
α_1	-0,02	0,01	-0,01	-0,01	0,01	0,01
σ_b	0,07	-0,05	0,05	-0,03	-0,09	0,03
σ_{u_1}	0,33	0,14	0,24	0,05	0,14	0,01
σ_{u_2}	2,69	1,08	1,06	0,40	0,81	0,29
ρ	-0,96	-0,77	-0,85	-0,71	-0,66	-0,47
ϕ	0,70	0,30	0,43	0,02	0,20	0,14
EQM						
β_0	0,97	0,10	0,21	0,04	0,12	0,03
β_1	0,80	0,13	0,29	0,07	0,13	0,05
α_0	0,02	0,00	0,01	0,00	0,01	0,00
α_1	0,02	0,00	0,01	0,00	0,01	0,00
σ_b	0,01	0,01	0,01	0,00	0,01	0,00
σ_{u_1}	0,00	0,00	0,00	0,00	0,00	0,00
σ_{u_2}	0,78	0,12	0,13	0,02	0,09	0,02
ρ	0,37	0,27	0,34	0,30	0,27	0,17
ϕ	17,22	14,42	16,16	0,08	12,08	2,20
PC						
β_0	0,94	0,98	0,96	0,98	0,98	0,94
β_1	0,94	0,98	0,96	0,90	0,94	0,92
α_0	0,92	0,98	0,96	0,90	0,94	0,94
α_1	0,90	0,98	0,96	0,88	1,00	0,96
σ_b	0,98	0,90	0,94	0,94	0,94	0,94
σ_{u_1}	1,00	1,00	0,98	0,98	0,98	0,98
σ_{u_2}	0,82	0,96	0,98	0,98	0,92	0,98
ρ	1,00	1,00	0,98	1,00	1,00	0,98
ϕ	1,00	0,90	0,90	0,96	0,88	0,96

Tabela B.5: Resultados da simulação segundo ER, EQM e PC para $\rho = -0,9$.

Parâmetro	Visitas = 5		Visitas = 10		Visitas = 15	
	$n = 50$	$n = 150$	$n = 50$	$n = 150$	$n = 50$	$n = 150$
ER						
β_0	0,26	0,10	0,14	0,04	0,05	0,04
β_1	0,19	-0,04	-0,09	0,02	0,03	0,01
α_0	0,00	-0,01	-0,01	-0,01	-0,01	0,00
α_1	0,01	0,01	0,02	0,02	0,02	0,00
σ	-0,02	-0,04	0,02	-0,02	-0,07	-0,02
σ_{u_1}	0,36	0,18	0,18	0,04	0,10	0,01
σ_{u_2}	2,76	0,96	1,04	0,63	1,02	0,31
ρ	-0,85	-0,81	-0,88	-0,57	-0,62	-0,45
ϕ	0,39	0,25	0,22	0,07	0,19	0,03
EQM						
β_0	1,15	0,14	0,27	0,04	0,20	0,04
β_1	0,81	0,11	0,23	0,06	0,27	0,04
α_0	0,01	0,00	0,01	0,00	0,00	0,00
α_1	0,02	0,00	0,01	0,00	0,01	0,00
σ	0,01	0,01	0,01	0,00	0,01	0,00
σ_{u_1}	0,00	0,00	0,00	0,00	0,00	0,00
σ_{u_2}	0,91	0,11	0,12	0,06	0,14	0,02
ρ	0,72	0,66	0,78	0,36	0,47	0,23
ϕ	6,73	7,16	3,02	1,14	8,38	0,33
PC						
β_0	0,90	0,90	0,94	0,98	0,92	0,92
β_1	0,94	0,96	0,98	0,96	0,98	0,96
α_0	0,98	0,96	0,94	0,90	0,98	0,96
α_1	0,96	0,98	0,90	0,90	0,96	0,96
σ	1,00	0,90	1,00	0,96	0,92	0,86
σ_{u_1}	1,00	1,00	1,00	1,00	1,00	1,00
σ_{u_2}	0,73	0,94	0,96	0,88	0,88	0,94
ρ	0,98	0,92	0,90	0,94	0,94	0,96
ϕ	0,94	0,92	0,98	0,94	0,96	0,94

Apêndice C

Análise da convergência

A Tabela C.1 apresenta os valores da estatística \hat{R} para os parâmetros do modelo final utilizando as distribuições Gama, Inversa Gaussiana e Lognormal.

Tabela C.1: Valores \hat{R} para os parâmetros do modelo final utilizando as distribuições Gama, Inversa Gaussiana e Lognormal.

Parâmetro	Gama	Inversa Gaussiana	Lognormal
β_1	1,00	1,00	1,00
β_2	1,00	1,00	1,00
β_3	1,00	1,00	1,00
β_4	1,00	1,00	1,00
β_5	1,00	1,00	1,00
β_6	1,00	1,00	1,00
β_7	1,00	1,00	1,00
α_1	1,00	1,00	1,02
α_2	1,00	1,00	1,00
α_3	1,01	1,00	1,01
α_4	1,01	1,00	1,01
α_5	1,00	1,00	1,01
σ_b	1,06	1,01	1,12
σ_{u_1}	1,00	1,00	1,01
σ_{u_2}	1,00	1,00	1,00
ρ	1,00	1,01	1,00
ϕ	1,00	1,00	1,16

As Figuras C.1, C.2 e C.3 apresentam gráficos do tipo Trace plot das cadeias para os parâmetros do modelo final escolhido utilizando a distribuição Gama.

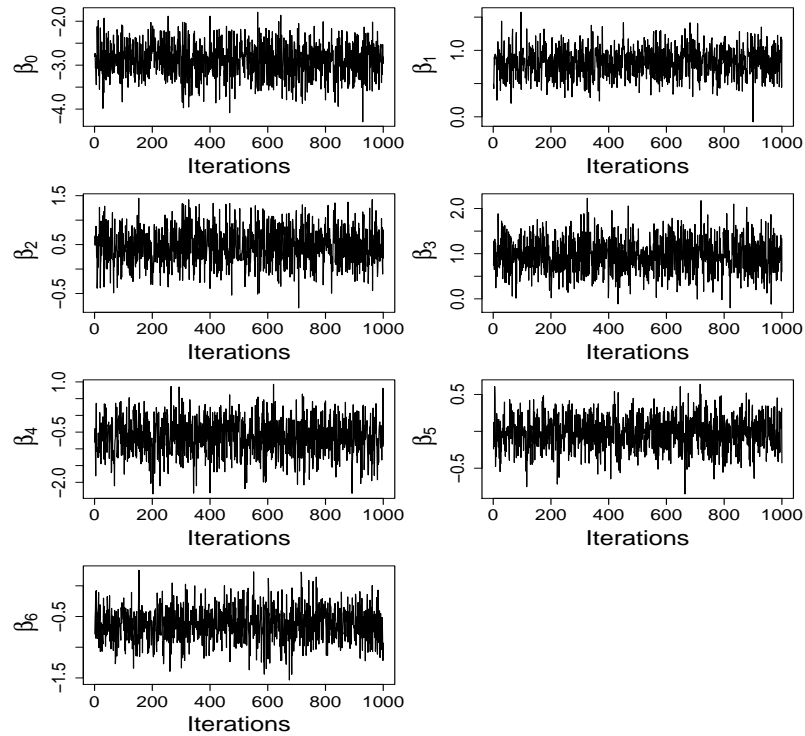


Figura C.1: Gráfico do tipo Trace plot para os parâmetros associados aos efeitos fixos da probabilidade de zero do modelo final utilizando a distribuição Gama.

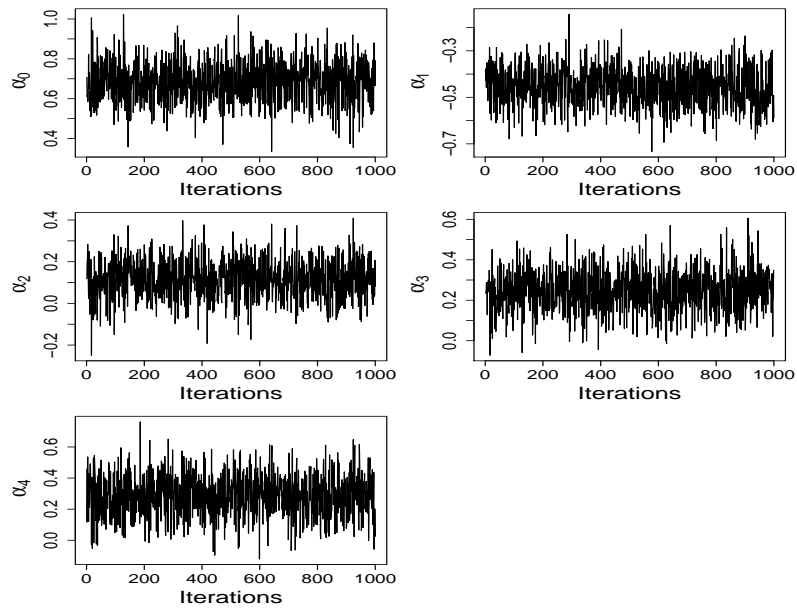


Figura C.2: Gráfico do tipo Trace plot para os parâmetros associados aos efeitos fixos da média dos valores positivos do modelo final utilizando a distribuição Gama.

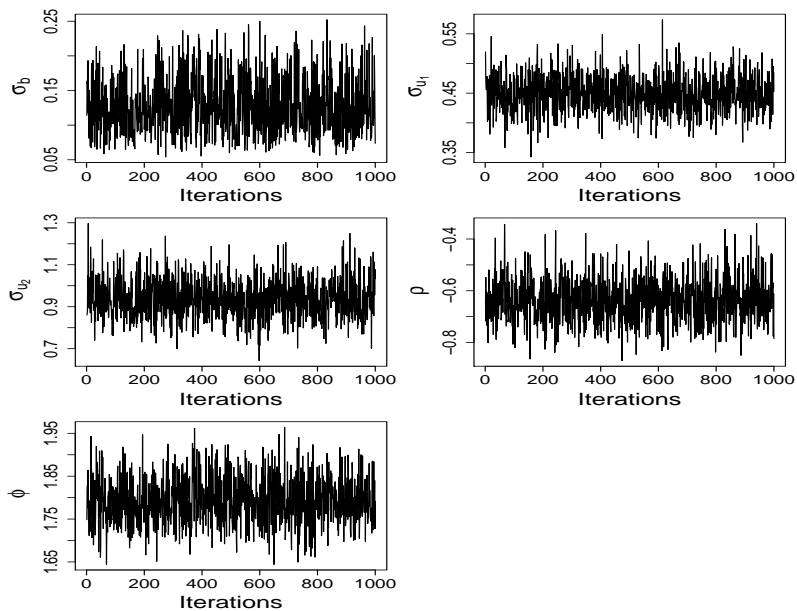


Figura C.3: Gráfico do tipo Trace plot para os parâmetros associados aos efeitos fixos das componentes de variância do modelo final utilizando a distribuição Gama.